

ADIABATIC REDUCTION OF A PIECEWISE DETERMINISTIC MARKOV MODEL OF STOCHASTIC GENE EXPRESSION WITH BURSTING TRANSCRIPTION

BY JINZHI LEI , MICHAEL C. MACKEY , ROMAIN YVINEC* AND CHANGJING ZHUGE

Tsinghua University, McGill University, Université de Lyon

This paper considers adiabatic reduction in a model of stochastic gene expression with bursting transcription. We prove that an adiabatic reduction can be performed in a stochastic slow/fast system with a jump Markov process. In the gene expression model, the production of mRNA (the fast variable) is assumed to follow a compound Poisson process (the phenomena called bursting in molecular biology) and the production of protein (the slow variable) is linear as a function of mRNA. When the dynamics of mRNA is assumed to be faster than the protein dynamics (due to a mRNA degradation rate larger than for the protein) we prove that, with the appropriate scaling, the bursting phenomena can be transmitted to the slow variable. We show that the reduced equation is either a stochastic differential equation with a jump Markov process or a deterministic ordinary differential equation depending on the scaling that is appropriate. These results are significant because adiabatic reduction techniques seem to have not been developed for a stochastic differential system containing a jump Markov process. We hope that they are generalizable to give insight into adiabatic methods in more general stochastic hybrid systems. In particular, a similar proof gives us a mathematical justification of the bursting of mRNA through an appropriate scaling of an “ON-OFF” switching gene model. Last but not least, for our particular system, the adiabatic reduction allows us to understand what are the necessary conditions for the bursting production-like of mRNA and protein to occur.

1. Introduction. The adiabatic reduction technique gives results that allow one to reduce the dimension of a system and justifies the use of an effective set of reduced equations in lieu of dealing with a full, higher dimensional, model if different time scales occur in the system. Adiabatic reduction results for deterministic systems of ordinary differential equations have

*corresponding author

AMS 2000 subject classifications: Primary 92C45, 60Fxx; secondary 92C40, 60J25, 60J75

Keywords and phrases: adiabatic reduction, piecewise deterministic Markov process, stochastic bursting gene expression, quasi-steady state assumption, scaling limit

been available since the very precise results of [Tikhonov \(1952\)](#) and [Fenichel \(1979\)](#). The simplest results, in the hyperbolic case, give an effective construction of an uniformly asymptotically stable slow manifold (and hence a reduced equation) and prove the existence of an invariant manifold near the slow manifold, with (theoretically) any order of approximation of this invariant manifold. Such precise and geometric results have been generalized to random systems of stochastic differential equation with Gaussian white noise ([Berglund and Gentz \(2006\)](#), see also [Gardiner \(1985\)](#) for previous work on the Fokker-Planck equation). However, to the best of our knowledge, analogous results for stochastic differential equations with jumps have not been obtained.

The present paper gives a theoretical justification of an adiabatic reduction of a particular piecewise deterministic Markov process ([Davis, 1984](#)). The results we obtain do not give a bound on the error of the reduced system, but they do allow us to justify the use of a reduced system in the case of a piecewise deterministic Markov process. In that sense, the results are close to the recent ones by [Crudu et al. \(2012\)](#) and [Kang and Kurtz \(2012\)](#), where general convergence results for discrete models of stochastic reaction networks are given. In particular, these papers give alternative scaling of the traditional ordinary differential equation and the diffusion approximation depending on the different scaling chosen (see [Ball et al. \(2006\)](#) for some examples in a reaction network model). After the scaling, the limiting models can be deterministic (ordinary differential equation), stochastic (jump Markov process), or hybrid (piecewise deterministic process). For illustrative and motivating examples given by a simulation algorithm, see [Haseltine and Rawlings \(2002\)](#); [Rao and Arkin \(2003\)](#); [Goutsias \(2005\)](#).

Our particular model is meant to describe stochastic gene expression with explicit bursting ([Friedman, Cai and Xie, 2006](#)). The variables evolve under the action of a continuous deterministic dynamical system interrupted by positive jumps of random sizes that model the burst production. In that sense, the convergence theorems we obtain in this paper can be seen as an example in which there are an *infinite* number of reactions (one for each “burst size”) and give complementary results to those of [Crudu et al. \(2012\)](#) and [Kang and Kurtz \(2012\)](#). We hope that the results here are generalizable to give insight into adiabatic reduction methods in more general stochastic hybrid systems ([Hespanha, 2006](#); [Bujorianu and Lygeros, 2004](#)). We note also that more geometrical approaches have been proposed to reduce the dimension of such systems in [Bujorianu and Katoen \(2008\)](#).

Biologically, the bursting of mRNA or protein is defined as the production of several molecules within a very short time, *indistinguishable* within

the time scale of the measurement. In the biological context of models of stochastic gene expression, explicit models of bursting mRNA and/or protein production have been analyzed recently, either using a discrete (Shahrezaei and Swain, 2008; Lei, 2009) or a continuous formalism (Friedman, Cai and Xie, 2006; Mackey, Tyran-Kaminska and Yvinec, 2011) as more and more experimental evidence from single-molecule visualization techniques has revealed the ubiquitous nature of this phenomenon (Ozbudak et al., 2002; Golding et al., 2005; Raj et al., 2006; Elf, Li and Xie, 2007; Xie et al., 2008; Raj and van Oudenaarden, 2009; Suter et al., 2011). Traditional models of gene expression are composed of *at least* two variables (mRNA and protein, and sometimes the DNA state). The use of a reduced one-dimensional model (that has the advantage that it can be solved analytically) has been justified so far by an argument concerning the stationary distribution in Shahrezaei and Swain (2008). However, it is clear that two different models may have the same stationary distribution but very different behavior (for an example in that context, see Mackey, Tyran-Kaminska and Yvinec (2011)). Hence, our results are of importance to rigorously prove the validity of using a reduced model.

1.1. *Continuous-state bursting model.* The models referred to above have explicitly assumed the production of several molecules *instantaneously*, through a jump Markov process, in agreement with experimental observations. In line with experimental observations, it is standard to assume a Markovian hypothesis (an exponential waiting time between production jumps) and that the jump sizes are exponentially distributed (geometrically in the discrete case) as well. The intensity of the jumps can be a bounded function to allow for self-regulation. The assumption that the intensity is bounded is crucial, and makes the stochastic production term a compound Poisson process.

Let X and Y denote the concentrations of mRNA and protein respectively. A simple model of single gene expression with bursting in transcription is given by

$$\begin{aligned} (1) \quad \frac{dX}{dt} &= -\gamma_1 X + \dot{N}(h, \varphi(Y)), \\ (2) \quad \frac{dY}{dt} &= -\gamma_2 Y + \lambda_2 X. \end{aligned}$$

Here γ_1 and γ_2 are the degradation rates for the mRNA and protein respectively, λ_2 is the mRNA translation rate, and $\dot{N}(h, \varphi(Y))$ describes the transcription that is assumed to be a compound Poisson *white noise* occurring at a rate φ with a non-negative jump size ΔX distributed with density

h .

The equations (1)-(2) are a short hand notation for

$$(3) \quad X_t = X_0 - \int_0^t \gamma_1 X_{s-} ds + \int_0^t \int_0^\infty \int_0^\infty 1_{\{r \leq \varphi(Y_{s-})\}} z N(ds, dz, dr)$$

$$(4) \quad Y_t = Y_0 - \int_0^t \gamma_2 Y_{s-} ds + \int_0^t \lambda_2 X_s ds.$$

where $X_{s-} = \lim_{t \rightarrow s-} X(t)$, and $N(ds, dz, dr)$ is a Poisson random measure on $(0, \infty) \times [0, \infty)^2$ with intensity $ds h(z) dz dr$, where s denotes the times of the jumps, r is the state-dependency in an acceptance/rejection fashion, and z the jump size. Note that (X_t) is a stochastic process with almost surely finite variation on any bounded interval $(0, T)$, so that the last integral is well defined as a Stieltjes-integral.

The following discussion is valid for general rate functions φ and density functions $h(\Delta X)$ that satisfy

$$(5) \quad \varphi \in C^1, \quad \varphi \text{ and } \varphi' \text{ are bounded, ie } \underline{\varphi} \leq \varphi, \varphi' \leq \bar{\varphi}$$

$$(6) \quad h \in C^0 \quad \text{and} \quad \int_0^\infty x^n h(x) dx < \infty, \quad \forall n \geq 1.$$

For a general density function h , the average burst size is given by

$$(7) \quad b = \int_0^\infty x h(x) dx.$$

REMARK 1. Hill functions are often used to model gene self-regulation, so that φ is given by

$$\varphi(y) = \kappa \frac{1 + y^n}{\Lambda + \Delta y^n}$$

where κ, Λ, Δ are positive parameters and n is a positive integer (see [Mackey, Tyran-Kaminska and Yvinec \(2011\)](#) for more details). An exponential distribution of the bursting transcription is often used in modeling gene expression, in accordance with experimental findings, so that the density function h is given by

$$h(\Delta X) = \frac{1}{b} e^{-\Delta X/b},$$

with b the average burst size.

If φ is independent of the state Y , the average transcription rate is $\lambda_1 =$

$b\varphi$, and the average mRNA and protein concentrations are

$$(8) \quad X_{\text{eq}} = \frac{b\varphi}{\gamma_1}.$$

$$(9) \quad Y_{\text{eq}} = \frac{\lambda_2}{\gamma_2} X_{\text{eq}} = \frac{b\varphi\lambda_2}{\gamma_1\gamma_2}.$$

1.2. *Statement of the results.* In the following discussion, we consider the situation when mRNA degradation is a fast process, *i.e.* γ_1 is ‘large enough’, but the average protein concentration Y_{eq} remains unchanged. Thus, we always assume one of the following three scaling relations:

- (S1) φ/γ_1 is independent of γ_1 ($= Y_{\text{eq}} \frac{\gamma_2}{b\lambda_2}$ if φ is a constant) as $\gamma_1 \rightarrow \infty$ and h, λ_2 remain unchanged; or
- (S2) $h \rightarrow \frac{1}{\gamma_1} h(\frac{\Delta X}{\gamma_1})$ (so that $b = Y_{\text{eq}} \frac{\gamma_2\gamma_1}{\varphi(Y)\lambda_2}$) as $\gamma_1 \rightarrow \infty$ and φ, λ_2 remain unchanged; or
- (S3) λ_2/γ_1 is independent of γ_1 ($= Y_{\text{eq}} \frac{\gamma_2\gamma_1}{b\varphi}$ if φ is a constant) as $\gamma_1 \rightarrow \infty$ and h, φ remain unchanged.

In this paper we determine an effective reduced equation for equation (2) for each of the three scaling conditions (S1)-(S3). In particular, we show that under assumption (S1), equation (2) can be approximated by the deterministic ordinary differential equation

$$(10) \quad \frac{dY}{dt} = -\gamma_2 Y + \lambda_2 \psi(Y)$$

where

$$(11) \quad \psi(Y) = b\varphi(Y)/\gamma_1.$$

We further show that under the scaling relations (S2) or (S3), equation (2) can be reduced to the stochastic differential equation

$$(12) \quad \frac{dY}{dt} = -\gamma_2 Y + \dot{N}(\bar{h}(\Delta Y), \varphi(Y)).$$

where \bar{h} is a suitable density function in the jump size ΔY (to be detailed below).

We first explain, using some heuristic arguments, the differences between the three scaling relations and the associated results. When $\gamma_1 \rightarrow \infty$, applying a standard quasi-equilibrium assumption we have

$$\frac{dX}{dt} \approx 0$$

which yields

$$(13) \quad X(t) \approx \frac{1}{\gamma_1} \dot{N}(h(\cdot), \varphi(y)) = \dot{N}(\gamma_1 h(\gamma_1 \cdot), \varphi(y)),$$

and therefore the second equation (2) becomes

$$\begin{aligned} \frac{dY}{dt} &\approx -\gamma_2 Y + \frac{\lambda_2}{\gamma_1} \dot{N}(h(\cdot), \varphi(y)) \\ &\approx -\gamma_2 Y + \dot{N}\left(\frac{\gamma_1}{\lambda_2} h\left(\frac{\gamma_1}{\lambda_2}\right), \varphi(y)\right). \end{aligned}$$

Hence in (12), $\bar{h}(\Delta Y) = (\lambda_2/\gamma_1)^{-1} h((\lambda_2/\gamma_1)^{-1} \Delta Y)$ under the scaling (S2) and (S3). Furthermore, we note that the scaling (S2) also implies $\gamma_1 h(\gamma_1 \cdot) = h(\cdot)$, and therefore we can also write $\bar{h}(\Delta Y) = (\lambda_2)^{-1} h((\lambda_2)^{-1} \Delta Y)$ under the scaling (S2). Finally, it is intuitively understandable that scaling (S1) gives (10), as the jumps become more frequent and smaller.

In this paper, we provide three different proofs of the results mentioned above. In particular, we prove the results using a Master equation approach (the Kolmogorov forward equation) as well as starting from the stochastic differential equation. Note that both techniques have been used in the past, in particular within the context of discrete models of stochastic reaction networks. For the Master equation approach, see [Haseltine and Rawlings \(2005\)](#); [Zeron and Santillán \(2010\)](#) while for the stochastic differential equation approach, we refer to [Crudu et al. \(2012\)](#); [Kang and Kurtz \(2012\)](#).

1.3. Outline of the paper. This paper is organized as follows. In Section 2, we consider the situation without auto-regulation so the rate φ is independent of protein concentration Y . In this case the two equations (1)-(2) form a set of linear stochastic differential equations. We use the method of characteristic functionals to give a very short proof of the adiabatic reduction when the white noise process is independent of the stochastic variables (external regulation). In Section 3 we consider the situation in which the rate φ is dependent on the protein concentration Y . We first give a result on the evolution equation on densities (Section 3.2), then we prove rigorously a weak convergence of the stochastic process using the machinery of generators (Section 3.3). Finally, in Section 4, we apply the same formalism with generators to prove that a binary “ON-OFF” model converges to a bursting model, thus illustrating that these techniques can be applicable in other contexts.

2. The case without auto-regulation. First, we consider the simple case with no auto-regulation in the mRNA transcription so the rate φ is independent of protein concentration Y . In this case, equation (1) is independent of (2) which makes the analysis easier.

Consider the equations

$$(14) \quad \frac{dX}{dt} = -\gamma_1 X + \dot{N}(h, \varphi), \quad X(0) = X_0 \geq 0,$$

$$(15) \quad \frac{dY}{dt} = -\gamma_2 Y + \lambda_2 X, \quad Y(0) = Y_0 \geq 0,$$

where $\dot{N}(h, \varphi)$ is a compound Poisson *white noise*, and φ is a constant. The solutions $X(t)$ and $Y(t)$ of (14)-(15) are stochastic processes uniquely determined by the equation parameters and the stochastic process \dot{N} .

We first state main results for each of the scaling relations we have considered.

REMARK 2. In all scaling situations that we consider in the three theorems below, the stochastic process X_t converges only in *finite-dimensional* law. Here *finite-dimensional* law means that for any finite number of fixed times (t_1, t_2, \dots, t_n) the random variable $(X(t_1), X(t_2), \dots, X(t_n))$ converges in law towards the appropriate distribution. Convergence in finite-dimensional law is weaker than convergence in law of the stochastic process. We will see in the proofs of the theorems that this is because a lack of compactness for the process $X(t)$.

THEOREM 1. Consider the equations (14)-(15). If the scaling (S1) is satisfied, i.e., $\varphi = Y_{\text{eq}} \frac{\gamma_2 \gamma_1}{b \lambda_2}$, then when $\gamma_1 \rightarrow \infty$,

1. The stochastic process $X(t)$ converges in finite-dimensional law to the (deterministic) steady-state value X_{eq} ;
2. The stochastic process $Y(t)$ converges in law towards the deterministic solution of the ordinary differential equation

$$(16) \quad \frac{dY}{dt} = -\gamma_2 Y + \lambda_2 X_{\text{eq}}, \quad Y(0) = Y_0 \geq 0.$$

THEOREM 2. Consider the equations (14)-(15). If the scaling (S2) is satisfied, i.e., $h \rightarrow \frac{1}{\gamma_1} h(\frac{\Delta X}{\gamma_1})$ (so $b = Y_{\text{eq}} \frac{\gamma_2 \gamma_1}{\varphi \lambda_2}$), then when $\gamma_1 \rightarrow \infty$,

1. The stochastic process $X(t)$ converges in finite-dimensional law to the compound Poisson white noise $\dot{N}(h, \varphi)$;

2. The stochastic process $Y(t)$ converges in law to the stochastic process defined by the solution of the stochastic differential equation

$$(17) \quad \frac{dY}{dt} = -\gamma_2 Y + \dot{N}(\bar{h}, \varphi), \quad Y(0) = Y_0 \geq 0$$

where $\bar{h}(\Delta Y) = (\lambda_2)^{-1} h((\lambda_2)^{-1} \Delta Y)$ is the density function for the jump size ΔY .

THEOREM 3. Consider the equations (14)-(15). If the scaling (S3) is satisfied, i.e., $\lambda_2 = Y_{\text{eq}} \frac{\gamma_2 \gamma_1}{b\varphi}$, then when $\gamma_1 \rightarrow \infty$,

1. The stochastic process $X(t)$ converges in finite-dimensional law to the (deterministic) fixed value 0;
2. The stochastic process $Y(t)$ converges in law to the stochastic process determined by the solution of the stochastic differential equation

$$(18) \quad \frac{dY}{dt} = -\gamma_2 Y + \dot{N}(\bar{h}, \varphi), \quad Y(0) = Y_0 \geq 0$$

where $\bar{h}(\Delta Y) = (\lambda_2/\gamma_1)^{-1} h((\lambda_2/\gamma_1)^{-1} \Delta Y)$ is the density function for the jump size ΔY .

REMARK 3. Note that scalings (S2) and (S3) give similar results for the equation governing the protein variable Y_t but very different results for the asymptotic stochastic process related to the mRNA. In particular, in Theorem 2, very large bursts of mRNA are transmitted to the protein, where in Theorem 3, very rarely is mRNA present but when present it is efficiently synthesized into a burst of protein.

These theorems will be proved using the method of characteristic functionals. For a stochastic process ξ_t ($t \geq 0$), the characteristic functional $C_\xi : \Sigma \rightarrow \mathbb{R}$ is defined as

$$(19) \quad C_\xi[f] = \mathbb{E} \left[e^{\int_0^\infty i f(t) \xi_t dt} \right]$$

for any function f in a suitable function space Σ so that the integral $\int_0^\infty i f(t) \xi_t dt$ is well defined. Before continuing, we need to introduce some topological background as well as properties of the Fourier transform in nuclear spaces (see Gel'fand and Vilenkin (1964))

2.1. *Stochastic process as a distribution.* We are going to recall here the continuous correspondence between a stochastic process and a distribution. All stochastic processes we consider here are right continuous with finite left limits. It is natural to define a stochastic process as a random element in the space $\mathcal{D} = \mathcal{D}((0, \infty), \mathcal{R})$, the space of all real-valued functions on $(0, \infty)$ that are right continuous with finite left limits. Such a space is classically endowed with the Skorokhod topology (see for instance [Jacod and Shiryaev \(1987, Chap. 6\)](#)). We define $D(\mathcal{R}^+)$, the space of smooth functions with compact support, with the inductive limit topology given by the family of semi-norms ($k = 0, 1, 2, \dots$) $p_k(f) = \sup |f^{(k)}|$ on every $D([0, n])$, $n \in \mathcal{N}$ (c.f. ([Schaefer, 1971](#), Example 2, page 57)). Let $f \in D(\mathcal{R}^+)$, and define \tilde{x} in the dual space $D'(\mathcal{R}^+)$ such that

$$(20) \quad \tilde{x}(f) = \int_0^\infty x(t)f(t)dt$$

LEMMA 4. *The map*

$$(21) \quad \begin{aligned} \mathcal{D}((0, \infty), \mathcal{R}) &\rightarrow D'(\mathcal{R}^+) \\ (x_t)_{t \geq 0} &\mapsto \tilde{x}, \end{aligned}$$

where \tilde{x} is defined by equation (20), is continuous.

PROOF. It is a classical result that $x \in \mathcal{D}$ has at most a countable number of discontinuity points so that x is locally integrable, the integral in Eq (20) is well defined for all $f \in D(\mathcal{R}^+)$, $\tilde{x} \in D'(\mathcal{R}^+)$ and

$$|\tilde{x}(f)| \leq \left(\int_0^T |x(s)| ds \right) \|f\|_\infty$$

for any f with support in $[0, T]$ ([Rudin, 1991](#), Section 6.11, page 142). We conclude by noticing that

$$|\tilde{x}(f)| \leq \sup_{s \leq T} |x(s)| \|f\|_\infty T$$

and $x \mapsto \sup_{s \leq T} |x(s)|$ is continuous for the Skorokhod topology ([Jacod and Shiryaev, 1987](#), Proposition 2.4, page 339) for all T such that T is not a discontinuity point. \square

2.2. *Bochner-Minlos theorem for a nuclear space.* Let E be a nuclear space. We state a key result that will allow us to uniquely identify a measure on the dual E' of E .

BOCHNER-MINLOS THEOREM. (*Gel'fand and Vilenkin, 1964, Theorem 2, page 146*) For a continuous functional C on a nuclear space E that satisfies $C(0) = 1$, and for any complex z_j and elements $x_j \in E$, $j, k = 1, \dots, n$,

$$(22) \quad \sum_{j=1}^n \sum_{k=1}^n z_j \bar{z}_k C(x_j - x_k) \geq 0,$$

there is a unique probability measure μ on the dual space E' , given by

$$(23) \quad C(y) = \int_{E'} e^{i\langle x, y \rangle} d\mu(x).$$

Note that the space $D(\mathcal{R}^+)$ is a nuclear space (*Schaefer, 1971, Example 2, page 107*).

2.3. *The characteristic functional of a Poisson white noise.* The use of the characteristic functional allows us to define a generalized stochastic process that does not necessarily have a trajectory in the usual sense (like in \mathcal{D} for instance). Indeed a (compound) Poisson white noise is seen as a random measure on the distribution space \mathcal{D} , associated with the characteristic functional (given in *Hida and Si (2004)*, here $f \in D(\mathcal{R}^+)$)

$$(24) \quad C_{\hat{N}}[f] = \exp \left[\varphi \int_0^\infty \int_0^\infty (e^{izf(t)} - 1) h(z) dz dt \right],$$

where φ is the Poisson intensity and h the jump size distribution. It is not hard to see that $C_{\hat{N}}[f - g]$ and $C_{\hat{N}}[g - f]$ are conjugate to each other, $C_{\hat{N}}[0] = 1$ and $C_{\hat{N}}[\cdot]$ is continuous for $h \in L^1(\mathcal{R}^+)$, so the conditions in the Bochner-Minlos Theorem 2.2 are satisfied and therefore $C_{\hat{N}}$ uniquely defines a measure on $D'(\mathcal{R}^+)$.

We refer to *Prokhorov (1956)*; *Hida and Si (2004, 2008)* for further material on characteristic functionals and generalized stochastic processes.

2.4. *Proofs of Theorems 1 through 3.* The proofs of Theorems 1 to 3 are based on the idea of Levy's continuity theorem. However in the infinite-dimensional case, the convergence of the Fourier transform does not imply convergence in law of the random variable, and one needs to impose more restrictions, namely a compactness condition. We will use the following lemma

LEMMA 5. *Let X^n be a sequence of stochastic processes in $\mathcal{D}((0, \infty), \mathcal{R})$. Suppose X^n is tight in $\mathcal{D}((0, \infty), \mathcal{R})$ and that there exists a random variable X such that, for all f in $D(\mathcal{R}^+)$, as $n \rightarrow \infty$,*

$$C_{X^n}[f] \rightarrow C_X[f].$$

Then X^n converges in law to X in $\mathcal{D}((0, \infty), \mathcal{R})$.

PROOF. The convergence of the characteristic functional, the Bochner-Minlos Theorem 2.2 and the continuity Lemma 4 ensure that the sequence X^n has at most one limiting law, which has to be the law of X . The classical Prokhorov Theorem (Jacod and Shiryaev, 1987, Corollary 3.9, page 348) states that tightness of X^n in $\mathcal{D}((0, \infty), \mathcal{R})$ is equivalent to relative compactness of the law of X^n in $\mathcal{P}(\mathcal{D})$, the space of probability measures on \mathcal{D} (with the topology of the weak convergence). Then X^n converges in law to X in $\mathcal{D}((0, \infty), \mathcal{R})$. \square

Now, we give the proofs of Theorems 1 through 3. The process is similar for each, and we only present a detailed proof for Theorem 1 and sketch the main differences in the proofs for Theorem 2 and 3.

Proof of Theorem 1. For any $f \in D(\mathcal{R}^+)$, from (14)-(15) and noting that the initial conditions X_0 and Y_0 are deterministic, it is not difficult to verify that (see also Cceres and Budini (1997))

$$(25) \quad G_X[f] = e^{ik_1 X_0} G_{\tilde{N}}[\tilde{f}_1(t)], \quad G_Y[f] = e^{ik_2 Y_0} G_X[\lambda_2 \tilde{f}_2(t)],$$

where

$$k_i = \int_0^\infty e^{-\gamma_i s} f(s) ds, \quad \tilde{f}_i(t) = \int_t^\infty e^{-\gamma_i(s-t)} f(s) ds, \quad (i = 1, 2).$$

Note that for any function $f \in D(\mathcal{R}^+)$ the functions $\tilde{f}_i(t)$ also belong to $D(\mathcal{R}^+)$ and therefore the characteristic functionals in (25) are well-defined. Furthermore, the characteristic functional of the compound Poisson white noise has been derived in Equation (24).

Now, we are ready to complete the proof by calculating the characteristic functionals G_X and G_Y when $\gamma_1 \rightarrow \infty$ from (25) and (24). First, we note that $k_i = \tilde{f}_i(0)$, and when $f \in D(\mathcal{R}^+)$ and $\gamma_1 \rightarrow \infty$,

$$(26) \quad \tilde{f}_1(t) = \frac{1}{\gamma_1} f(t) + O\left(\frac{1}{\gamma_1^2}\right).$$

Furthermore, from (25)-(26), we have

$$\begin{aligned}
 G_X[f] &= e^{iX_0(\frac{1}{\gamma_1}f(0)+O(\frac{1}{\gamma_1^2}))} G_N \left[\frac{1}{\gamma_1}f(t) + O(\frac{1}{\gamma_1^2}) \right] \\
 (27) \quad &= e^{i\frac{1}{\gamma_1}X_0f(0)} \exp \left[i \int_0^\infty \int_0^\infty \frac{\varphi}{\gamma_1} f(t) x h(x) dx dt \right] + O(\frac{1}{\gamma_1}).
 \end{aligned}$$

Thus, from the scaling (S1) and (7), we have

$$(28) \quad \lim_{\gamma_1 \rightarrow \infty} G_X[f] = \exp \left[i \int_0^\infty f(t) X_{eq} dt \right], \quad \forall f \in D.$$

Therefore, (25) yields

$$\begin{aligned}
 \lim_{\gamma_1 \rightarrow \infty} G_Y[f] &= e^{ik_2 Y_0} \exp \left[i \lambda_2 \int_0^\infty \tilde{f}_2(t) X_{eq} dt \right] \\
 (29) \quad &= e^{ik_2 Y_0} \exp \left[i \int_0^\infty f(s) (1 - e^{-\gamma_2 s}) Y_{eq} ds \right].
 \end{aligned}$$

Now, it is easy to verify that the right hand sides of (28) and (29) give, respectively, the characteristic functional of $X(t) \equiv X_{eq}$ and $Y(t)$ of the solution of (16). Thus, the process $X(t)$ of (14) converges in finite-dimensional law to X_{eq} (by Levy's continuity theorem).

Finally, to show that the process $Y(t)$ defined by (15) converges in law towards the solution of (16), we apply Lemma 5 and only need to show that a sequence $\{Y_t^n\}$ defined by $\{\lambda_{1,n}\}$ with $\lambda_{1,n} \rightarrow \infty$ is tight in $\mathcal{D}((0, \infty), \mathcal{R})$. Let $\{\lambda_{1,n}\}$ be a sequence of positive constants so that $\lambda_{1,n} \rightarrow \infty$ when $n \rightarrow \infty$, and $\{X_t^n\}$ and $\{Y_t^n\}$ are two sequences of stochastic processes such that (X_t^n, Y_t^n) is determined by (14)-(15) with λ_1 replaced by $\lambda_{1,n}$.

For any n , let N_n be a compound Poisson process with $\{T_{n,i}\}_{i=1}^\infty$ the jump times which occur at a rate $\varphi_n = Y_{eq} \frac{\gamma_2 \gamma_{1,n}}{b \lambda_2}$, and $\{Z_{n,i}\}_{i=1}^\infty$ the jump sizes that are iid random variables with density h (with the convention $T_{n,0} = 0$ and $Z_{n,0} = X_0$). Then

$$X_t^n = \sum_{T_{n,i} \leq t} Z_{n,i} e^{-\gamma_1(t-T_{n,i})} 1_{t \geq T_{n,i}}.$$

From (4), we have

$$Y_t^n \leq Y_0 + \int_0^t \lambda_2 X(s) ds \leq Y_0 + \frac{\lambda_2}{\gamma_{1,n}} \sum_{T_{n,i} \leq t} Z_{n,i} = Y_0 + \frac{\lambda_2 N_n(t)}{\gamma_{1,n}}.$$

Now,

$$N_n(t) = \tilde{N}_n(t) + \varphi_n t \int_0^\infty zh(z)dz = \tilde{N}_n(t) + \varphi_n bt,$$

where \tilde{N}_n is a *compensated* compound Poisson process (that is $\mathbb{E}[\tilde{N}_n] = 0$). \tilde{N}_n is a right continuous Martingale. Then by Doob's inequalities ([Ethier and Kurtz, 2005](#), Corollary 2.17, page 64)),

$$\mathcal{P}\left(\sup_{t \leq T} |\tilde{N}_n(t)| \geq K\right) \leq \frac{1}{K} \mathbb{E} |\tilde{N}_n(T)|.$$

Now let $T > 0$, $\varepsilon > 0$, and K be such that $\mathcal{P}\left(|Y_0| \geq K/2\right) \leq \frac{\varepsilon}{2}$, so we have

$$(30) \quad \mathcal{P}\left(\sup_{t \leq K} |Y_t^n| \geq K\right) \leq \frac{\varepsilon}{2} + \frac{4}{K} \frac{\lambda_2 \varphi_n b T}{\gamma_{1,n}} \leq \varepsilon$$

for K sufficiently large, according to hypothesis (6) and the scaling (S1). Similarly, for any $t_1, t_2 \in [0, T]$,

$$|Y_{t_2}^n - Y_{t_1}^n| \leq \frac{\lambda_2}{\gamma_{1,n}} |N_n(t_2) - N_n(t_1)|$$

and, for any $T > 0$, $\varepsilon > 0$, $\eta > 0$,

$$(31) \quad \mathcal{P}\left(|Y_{t_2}^n - Y_{t_1}^n| \geq \eta\right) \leq \frac{2\lambda_2 \varphi_n b |t_2 - t_1|}{\gamma_{1,n} \eta} \leq \varepsilon$$

for $|t_2 - t_1|$ sufficiently small. With (30)-(31) and by classical criteria for tightness in \mathcal{D} ([Jacod and Shiryaev, 1987](#), Theorem 3.21, page 350) $\{Y_t^n\}$ is tight in $\mathcal{D}((0, \infty), \mathcal{R})$.

□

Proof of Theorem 2. Let

$$h_{\gamma_1}(x) = \frac{1}{\gamma_1} h\left(\frac{x}{\gamma_1}\right).$$

The proof is similar to the proof of Theorem 2. Note simply from the scaling (S2) that (27) becomes

$$(32) \quad G_X[f] = e^{i\frac{1}{\gamma_1} X_0 f(0)} \exp \left[i\varphi \int_0^\infty \int_0^\infty (e^{if(t)x/\gamma_1} - 1) h_{\gamma_1}(x) dx dt \right] + O\left(\frac{1}{\gamma_1}\right).$$

Thus, we have

$$\begin{aligned}
\lim_{\gamma_1 \rightarrow \infty} G_X[f] &= \lim_{\gamma_1 \rightarrow \infty} \exp \left[i\varphi \int_0^\infty \int_0^\infty (e^{if(t)x/\gamma_1} - 1) h\left(\frac{x}{\gamma_1}\right) d\left(\frac{x}{\gamma_1}\right) dt \right] \\
&= \exp \left[i\varphi \int_0^\infty \int_0^\infty (e^{if(t)x} - 1) h(x) dx dt \right] \\
&= G_{\hat{N}}[f].
\end{aligned}$$

Furthermore, from (25)

$$\begin{aligned}
\lim_{\gamma_1 \rightarrow \infty} G_Y[f] &= e^{ik_2 Y_0} G_{\hat{N}}[\lambda_2 \tilde{f}_2(t)] \\
&= e^{ik_2 Y_0} \exp \left[i\varphi \int_0^\infty \int_0^\infty (e^{i\lambda_2 x \tilde{f}_2(t)} - 1) h(x) dx dt \right] \\
(33) \quad &= e^{ik_2 Y_0} \exp \left[i\varphi \int_0^\infty \int_0^\infty (e^{ix \tilde{f}_2(t)} - 1) \bar{h}(x) dx dt \right],
\end{aligned}$$

where

$$\bar{h}(x) = \lambda_2^{-1} h(\lambda_2^{-1} x).$$

It is easy to verify that (33) is just the characteristic functional of the stochastic processes given by solutions of (17).

The rest of the proof is similar to the proof of Theorem 1 (Equations (30)-(31) are still valid) and the details are omitted.

□

Proof of Theorem 3 Here, it is sufficient to note that $\lambda_2 = Y_{\text{eq}} \frac{\gamma_2}{\gamma_1} b\varphi$ in the scaling (S3), and furthermore

$$(34) \quad \lim_{\gamma_1 \rightarrow \infty} G_X[f] = \lim_{\gamma_1 \rightarrow \infty} \exp \left[\varphi \int_0^\infty \int_0^\infty (e^{if(t)x/\gamma_1} - 1) h(x) dx dt \right] = 1,$$

and

$$\begin{aligned}
\lim_{\gamma_1 \rightarrow \infty} G_Y[f] &= e^{ik_2 Y_0} \exp \left[\varphi \int_0^\infty \int_0^\infty (e^{i(\lambda_2/\gamma_1)x \tilde{f}_2(t)} - 1) h(x) dx dt \right] \\
(35) \quad &= e^{ik_2 Y_0} \exp \left[\varphi \int_0^\infty \int_0^\infty (e^{ix \tilde{f}_2(t)} - 1) \bar{h}(x) dx dt \right],
\end{aligned}$$

where

$$\bar{h}(x) = \frac{\gamma_1}{\lambda_2} h\left(\frac{\gamma_1}{\lambda_2} x\right).$$

□

3. The case with auto-regulation. We now turn to a consideration of the adiabatic reduction of equations (1)-(2) in the general case when the jump rate φ is dependent on protein concentration, $\varphi(Y)$.

We use two different methods to arrive at our results. The first method is based on the density evolution equations, and shows that the evolution equations obtained from the two stochastic differential equations are consistent with each other when $\gamma_1 \rightarrow +\infty$. Existence of densities for such processes has been rigorously proved in [Tyran-Kaminska \(2009\)](#); [Mackey and Tyran-Kamińska \(2008\)](#). The second method is based on the convergence of generators and shows that the stochastic process $Y(t)$ given by equations (1)-(2) converges in distribution to that of the solution of equation (12). Background on generators and convergence of stochastic processes can be found in [Ethier and Kurtz \(2005\)](#), and generators for hybrid systems have been characterized in [Davis \(1984\)](#); [Hespanha \(2006\)](#); [Bujorianu and Lygeros \(2004\)](#). Though the second method implies the first, the two proofs using the two approaches are interesting and give different insights into the problem.

We first summarize the important background results on the stochastic processes described in our paper.

One dimensional equation. For the one-dimensional stochastic differential equation (12) perturbed by a compound Poisson white noise, of (bounded) intensity $\varphi(y)$ and jump size distribution \bar{h} , the infinitesimal generator of the stochastic process $(Y_t)_{t \geq 0}$ is ([Davis, 1984](#), Theorem 5.5), for any $f \in \mathcal{D}(\mathcal{A})$,

$$(36) \quad \mathcal{A}_1 f(y) = -\gamma_2 y \frac{df}{dy} + \varphi(y) \left(\int_y^\infty \bar{h}(z - y) f(z) dz - f(y) \right)$$

$$(37) \quad \mathcal{D}(\mathcal{A}_1) = \{f \in \mathcal{M}(0, \infty) : t \mapsto f(ye^{-\gamma_2 t}) \text{ is absolutely continuous for } t \in \mathcal{R}^+ \text{ and} \\ \mathbb{E} \sum_{T_i \leq t} |f(Y_{T_i}) - f(Y_{T_i}^-)| < \infty \text{ for all } t \geq 0\}$$

where $\mathcal{M}(0, \infty)$ denotes a Borel-measurable function of $(0, \infty)$ and the times T_i are the instants of the jump of y_t . It is an extended domain containing all functions that are sufficiently smooth along the deterministic trajectories between the jumps, and with a bounded total variation induced by the jumps.

The operator \mathcal{A}_1 is the adjoint of the operator acting on densities $v(t, y)$ given by [Mackey and Tyran-Kamińska \(2008\)](#)

$$(38) \quad \frac{\partial v(t, y)}{\partial t} = \frac{\partial}{\partial y} [\gamma_2 y v(t, y)] + \int_0^y \varphi(z) v(t, z) \bar{h}(y - z) dz - \varphi(y) v(t, y).$$

For any $f \in \mathcal{D}(\mathcal{A}_1)$, we have

$$(39) \quad \frac{d}{dt} \mathbb{E}f(X_t) = \mathbb{E}\mathcal{A}_1(f(X_t)).$$

Two dimensional equation. Consideration of the two-dimensional stochastic differential equation (1-2) perturbed by a compound Poisson white noise, of intensity $\varphi(y)$ and jump size distribution h follows along similar lines. Its infinitesimal generator and extended domain are

$$(40) \quad \begin{aligned} \mathcal{A}_2 g(x, y) = & -\gamma_1 x \frac{\partial g}{\partial x} + (\lambda_2 x - \gamma_2 y) \frac{\partial g}{\partial y} \\ & + \varphi(y) \left(\int_x^\infty h(z - x) g(z, y) dz - g(x, y) \right), \end{aligned}$$

$$(41) \quad \begin{aligned} \mathcal{D}(\mathcal{A}_2) = & \{g \in \mathcal{M}((0, \infty)^2) : t \mapsto g(\phi_t(x, y)) \text{ is absolutely} \\ & \text{continuous for } t \in \mathcal{R}^+ \text{ and} \\ & \mathbb{E} \sum_{T_i \leq t} |g(X_{T_i}, Y_{T_i}) - g(X_{T_i^-}, Y_{T_i^-})| < \infty \text{ for all } t \geq 0\} \end{aligned}$$

where ϕ_t is the deterministic flow given by equations (1-2).

The evolution equation for densities $u(t, x, y)$ is

$$(42) \quad \begin{aligned} \frac{\partial u(t, x, y)}{\partial t} = & \frac{\partial}{\partial x} [\gamma_1 x u(t, x, y)] - \frac{\partial}{\partial y} [(\lambda_2 x - \gamma_2 y) u(t, x, y)] \\ & + \int_0^x \varphi(y) u(t, z, y) h(x - z) dz - \varphi(y) u(t, x, y). \end{aligned}$$

For any $f \in \mathcal{D}(\mathcal{A}_2)$, we have

$$(43) \quad \frac{d}{dt} \mathbb{E}f(X_t, Y_t) = \mathbb{E}\mathcal{A}_2(f(X_t, Y_t)).$$

Note that from the results of Section 2, demonstrating convergence in law for the first variable X_t is impossible. Rather, the finite-dimensional convergence properties are replaced by scaling behavior of the marginal moment. As expected, this scaling depends on the hypotheses (S1), (S2) or (S3). This scaling behavior is crucial for the convergence results and is discussed in the following subsection before we prove the main results.

3.1. *Scaling of the marginal moment.* Using the generator \mathcal{A}_2 for the two-dimensional stochastic process defined by (41), we can deduce the scaling laws of the marginal moment of X_t^n as $\gamma_1^n \rightarrow \infty$.

PROPOSITION 6. *Let $(X(t), Y(t))$ be the solutions of (1)-(2), and $\mu_k(t) = \mathbb{E}[X(t)^k]$ and $\nu_k(t) = \mathbb{E}[Y(t)X(t)^k]$. Suppose $\mu_k(0) < \infty$ and $\nu_k(0) < \infty$, then $\mu_k(t) < \infty$ and $\nu_k(t) < \infty$ for all t . Moreover, for fixed $t > 0$,*

1. *If the scaling (S1) holds, then both $\mu_k(t)$ and $\nu_k(t)$ stay uniformly bounded above and below as $\gamma_1 \rightarrow \infty$.*
2. *For the scaling (S2), then, for $k \geq 1$,*

$$(44) \quad \mu_k(t) \sim \gamma_1^{k-1}, \quad \nu_k(t) \sim \gamma_1^{k-1}, \quad (\gamma_1 \rightarrow \infty)$$

and $\nu_0(t)$ is uniformly bounded above and below as $\gamma_1 \rightarrow \infty$.

3. *If (S3) holds then, for $k \geq 1$,*

$$(45) \quad \mu_k(t) \sim \gamma_1^{-1}, \quad \nu_k(t) \sim \gamma_1^{-1}, \quad (\gamma_1 \rightarrow \infty)$$

and $\nu_0(t)$ is uniformly bounded above and below as $\gamma_1 \rightarrow \infty$.

PROOF. The proposition is proved using the evolution equation for the marginal moment obtained from the generator \mathcal{A}_2 .

First, we claim that functions x^k and $x^k y$ ($\forall k \in \mathbb{N}^*$) are contained in $\mathcal{D}(\mathcal{A}_2)$. To show this, we only need to verify that

$$(46) \quad \mathbb{E} \sum_{T_i \leq t} |X(T_i)^k Y(T_i)^l - X(T_i^-)^k Y(T_i^-)^l| < \infty, \quad \forall t \geq 0, k \in \mathbb{N}^*, l = 0, 1,$$

where the T_i are jump times. Since $Y(t)$ is continuous and from equation (4), $\mathbb{E}Y_t \leq \mathbb{E}Y_0 + \int_0^t \mathbb{E}X_s ds$, we only need to verify the case with $l = 0$.

Let Z_i be the random variable associated with the i^{th} jump, and N_t the number of jumps for which $T_i \leq t$. Then $X(T_i) = X(T_i^-) + Z_i$, and thus

$$\begin{aligned} \mathbb{E} \sum_{T_i \leq t} |(X(T_i)^k - X(T_i^-)^k)| &= \mathbb{E} \sum_{T_i \leq t} \sum_{j < k} \binom{k}{j} X(T_i^-)^j Z_i^{k-j} \\ &\leq \mathbb{E} \sum_{i=1}^{N_t} \sum_{j < k} \binom{k}{j} \left(\sum_{l=1}^i Z_l \right)^j Z_i^{k-j} \\ &\leq (\mathbb{E}N_t) C(\mathbb{E}h, \mathbb{E}^2h, \dots, \mathbb{E}^k h) < \infty, \end{aligned}$$

where $\mathbb{E}^j h = \int_0^\infty x^j h(x) dx$ is the j^{th} moment of h , and $C(\mathbb{E}h, \mathbb{E}^2h, \dots, \mathbb{E}^k h)$ is a polynomial in the first k^{th} moments. Here we have used the independence

of the jump times and the jump sizes, as well as the upper bound of the trajectories X_t due to the following formal expression

$$X_t = \sum_{T_i \leq t} H(t - T_i) Z_i e^{-\gamma(t - T_i)}$$

where H stands for the Heaviside step function.

Now, $\mathcal{A}_2 x^k$ and $\mathcal{A}_2 x^k y$ are well defined. A straightforward calculation yields

$$\begin{aligned} \mathcal{A}_2 x^k &= -\gamma_1 k x^k + \varphi(y) \left(\int_x^\infty h(z - x)(z - x + x)^k dz - x^k \right) \\ &= -\gamma_1 k x^k + \varphi(y) \sum_{i=0}^{k-1} \binom{k}{i} x^i \int_x^\infty h(z - x)(z - x)^{k-i} dz \\ &= -\gamma_1 k x^k + \varphi(y) \sum_{i=0}^{k-1} \binom{k}{i} x^i \mathbb{E}^{k-i} h. \end{aligned}$$

Then the k^{th} -marginal moment $\mu_k(t)$ of the first variable x depends only on the lower moment $\mu_i(t)$, $i < k$. We then obtain, with assumption (5) and equation (43)

$$\begin{aligned} -\gamma_1 k \mu_k(t) + \underline{\varphi} \sum_{i=0}^{k-1} \binom{k}{i} \mu_i(t) \mathbb{E}^{k-i} h &\leq \dot{\mu}_k(t), \\ (47) \quad \dot{\mu}_k(t) &\leq -\gamma_1 k \mu_k(t) + \overline{\varphi} \sum_{i=0}^{k-1} \binom{k}{i} \mu_i(t) \mathbb{E}^{k-i} h. \end{aligned}$$

Assume scaling (S1). Applying Gronwall's inequality to equation (47) for $k = 1$ yields, for all $t > 0$,

$$\frac{\underline{\varphi} b}{\gamma_1} + O\left(\frac{1}{\gamma_1^2}\right) \leq \mu_1(t) \leq \frac{\overline{\varphi} b}{\gamma_1} + O\left(\frac{1}{\gamma_1^2}\right)$$

Iteratively, for all $t > 0$ and $k > 1$, there is a constant $c_k > 0$ independent of γ_1 (where c_k depends only on the moment of h) such that

$$\frac{\underline{\varphi} c_k}{\gamma_1} + O\left(\frac{1}{\gamma_1^2}\right) \leq \mu_k(t) \leq \frac{\overline{\varphi} c_k}{\gamma_1} + O\left(\frac{1}{\gamma_1^2}\right).$$

Assume (S2). The case $k = 1$ follows directly from the above calculations, and for all $k > 1$ and $t > 0$,

$$\frac{\underline{\varphi} \gamma_1^k \mathbb{E}_k h}{k \gamma_1} + O(\gamma_1^{k-2}) \leq \mu_k(t) \leq \frac{\overline{\varphi} \gamma_1^k \mathbb{E}_k h}{\gamma_1} + O(\gamma_1^{k-2}).$$

Finally, assume (S3). The same method shows that for all $t > 0$ and $k \geq 1$, there is a constant c_k independent of γ_1 (c_k depends of the moment of h and of φ) such that

$$\frac{c_k}{\gamma_1} + O\left(\frac{1}{\gamma_1^2}\right) \leq \mu_k(t) \leq \frac{c_k}{\gamma_1} + O\left(\frac{1}{\gamma_1^2}\right).$$

A similar calculation with $g(x, y) = yx^k$ gives analogous scaling. Namely, we have

$$\mathcal{A}_2 x^k y = (-\gamma_1 k + \gamma_2) x^k y + \lambda_2 x^{k+1} + \varphi(y) \sum_{i=0}^{k-1} \binom{k}{i} x^i \mathbb{E}^{k-i} h$$

so that, for $k \geq 1$,

$$\begin{aligned} & (-\gamma_1 k + \gamma_2) \nu_k(t) + \lambda_2 \mu_{k+1} + \varphi \sum_{i=0}^{k-1} \binom{k}{i} \mu_i(t) \mathbb{E}^{k-i} h \\ & \leq \dot{\nu}_k(t) \leq (-\gamma_1 k + \gamma_2) \nu_k(t) + \lambda_2 \mu_{k+1} + \bar{\varphi} \sum_{i=0}^{k-1} \binom{k}{i} \mu_i(t) \mathbb{E}^{k-i} h \end{aligned}$$

while for $k = 0$, we obtain

$$\dot{\nu}_0 = -\gamma_2 \nu_0 + \lambda_2 \mu_1.$$

Then ν_0 is uniformly bounded for each scaling (S1), (S2), and (S3). Then, iteratively using the inequalities for $\dot{\nu}_k$, the scaling of μ_{k+1} and Gronwall's inequality yields the desired result for each scaling. \square

3.2. Density evolution equations. Let $u(t, x, y)$ be the density function of (X, Y) at time t obtained from the solutions of equation (1)-(2). The evolution of the density $u(t, x, y)$ is governed by

$$\begin{aligned} (48) \quad \frac{\partial u(t, x, y)}{\partial t} &= \frac{\partial}{\partial x} [\gamma_1 x u(t, x, y)] - \frac{\partial}{\partial y} [(\lambda_2 x - \gamma_2 y) u(t, x, y)] \\ &\quad + \int_0^x \varphi(y) u(t, z, y) h(x - z) dz - \varphi(y) u(t, x, y) \end{aligned}$$

when $(t, x, y) \in (0, +\infty)^3$.

In this subsection, we prove that when $\gamma_1 \rightarrow \infty$ the density function $u(t, x, y)$ approaches the density $v(t, y)$ for solutions of either the deterministic equation (10) or the stochastic differential equation (12) depending on

the scaling. Evolution of the density function for equation (10) is given by [Lasota and Mackey \(1985\)](#)

$$(49) \quad \frac{\partial v(t, y)}{\partial t} = -\frac{\partial}{\partial y}[-\gamma_2 y u_0 + \lambda_2 \psi(y) u_0].$$

Here we note that

$$(50) \quad \psi(y) = b\varphi(y)/\gamma_1.$$

Evolution of the density for equation (12) is given by

$$(51) \quad \frac{\partial v(t, y)}{\partial t} = \frac{\partial}{\partial y}[\gamma_2 y v(t, y)] + \int_0^y \varphi(z) v(t, z) \bar{h}(y - z) dz - \varphi(y) v(t, y)$$

when $(t, y) \in (0, +\infty)^2$. Here \bar{h} is given by

$$(52) \quad \bar{h}(y) = \frac{\gamma_1}{\lambda_2} h\left(\frac{\gamma_1}{\lambda_2} y\right).$$

We note that under the scaling (S2), $\bar{h}(y) = \lambda_2^{-1} h(\lambda_2^{-1} y)$ is independent of γ_1 , and under the scaling (S3), γ_1/λ_2 is independent of γ_1 , and therefore \bar{h} is also independent of γ_1 .

In the following proof, we will need existence of derivatives of all order of the densities functions. Thus, in equations (48)-(51) we make the additional hypotheses:

- (H1) The density function h satisfies (6), and $h \in \mathcal{C}^\infty$;
- (H2) The function \bar{h} is given by Theorem 2 or Theorem 3 depending on the scaling;
- (H3) The rate function $\varphi(Y)$ satisfies (5) and $\varphi \in \mathcal{C}^\infty$.

When conditions (5)-(6) are satisfied, existence of the above densities have been rigorously proved in [Mackey and Tyran-Kamińska \(2008\)](#); [Tyran-Kamińska \(2009\)](#). In particular, for a given initial density

$$(53) \quad u(0, x, y) = p(x, y), \quad 0 < x, y < +\infty$$

that satisfies

$$(54) \quad p(x, y) \geq 0, \quad \int_0^\infty \int_0^\infty p(x, y) dx dy = 1,$$

there is a unique solution $u(t, x, y)$ of (48) that satisfies the initial condition (53) and

$$(55) \quad u(t, x, y) \geq 0, \quad \int_0^\infty \int_0^\infty u(t, x, y) dx dy = 1$$

Moreover, if the moments of the initial density satisfy

$$(56) \quad u_n(y) = \int_0^\infty x^n p(x, y) dx < +\infty, \quad \forall y > 0, n = 0, 1, \dots,$$

then the marginal moments

$$(57) \quad u_n(t, y) = \int_0^\infty x^n u(t, x, y) dx,$$

are well defined for $t > 0$ and $y > 0$ from the discussion in Section 3.1. Therefore

$$(58) \quad \lim_{x \rightarrow \infty} x^n u(t, x, y) = 0, \quad \forall t, y.$$

Here, we will show, using semigroup techniques as in [Mackey and Tyran-Kamińska \(2008\)](#); [Tyran-Kamińska \(2009\)](#), that under the hypotheses (H1)-(H3), the densities are smooth. We will use the following result

PROPOSITION 7. (*Pazy, 1992, Corollary 5.6, page 124*) *Let Y be a subspace of a Banach space X , with $(Y, \|\cdot\|_Y)$ a Banach space as well. Let $T(t)$ be a strongly continuous semigroup on X , with infinitesimal generator C . Then Y is an invariant subspace of $T(t)$ if*

- *For sufficiently large λ , Y is an invariant subspace of $R(\lambda, C)$*
- *There exist constants c_1 and c_2 such that, for $\lambda > c_2$,*

$$\|R(\lambda, C)^n\|_Y \leq c_1(\lambda - c_2)^{-n}, \quad n = 1, 2, \dots$$

- *For $\lambda > c_2$, $R(\lambda, C)Y$ is dense in Y .*

Then, we have

LEMMA 8. *Assume (H1)-(H3). If the initial condition $v(0, y) \in \mathcal{C}^\infty \cap L^1$ then the unique solution of equation (51) (respectively (49)) $v(t, y) \in \mathcal{C}^\infty \cap L^1$. Similarly if the initial condition $u(0, x, y) \in (\mathcal{C}^\infty)^2 \cap L^1$ then the unique solution of equation (48) $u(t, x, y) \in (\mathcal{C}^\infty)^2 \cap L^1$.*

PROOF. Because the dynamical system given by (10) is smooth and invertible, the result for equation (49) is standard ([Lasota and Mackey, 1985](#), Remark 7.6.2 page 187). We will show that the result for equation (51), and the result for equation (48) will follow in a similar fashion. We need to show that the subspace $\mathcal{C}_0^\infty \subseteq L^1$ is invariant under the action of the semigroup defined by equation (51). According to [Mackey and Tyran-Kamińska \(2008\)](#)

(and references therein), we know that the semigroup defined by equation (51) is a strongly continuous semigroup whose infinitesimal generator C is characterized by the resolvent

$$(59) \quad R(\lambda, C)v = \lim_{N \rightarrow \infty} R(\lambda, A) \sum_{n=0}^N (P(\varphi R(\lambda, A)))^n v$$

for all $v \in L^1$, $\lambda > 0$, where the limit holds in L^1 and A and P are the operators given by

$$(60) \quad Av(y) = \frac{d(\gamma_2 y v)}{dy} - \varphi(y)v(y)$$

$$(61) \quad Pv(y) = \int_0^y v(z)h(y-z)dz,$$

and the resolvent $R(\lambda, A)$ is given by, for all $v \in L^1$,

$$(62) \quad R(\lambda, A)v(y) = \int_y^\infty \frac{1}{\gamma_2 y} e^{Q_\lambda(z) - Q_\lambda(y)} v(z) dz$$

with $Q_\lambda(y) = -\frac{\lambda \ln(y)}{\gamma_2} - \int_1^y \frac{\varphi(z)}{\gamma_2 z} dz$. We also know that for

$$v \in \mathcal{D}(A) = \{v \in L^1 : (yv) \text{ is absolutely continuous and } \left(\frac{d(yv)}{dy}\right) \in L^1\},$$

we have

$$(63) \quad Cv = Av + P(\varphi v).$$

We will now use the result from by Proposition 7 above to complete the proof. Note that according to (H1)-(H3), Q_λ is a \mathcal{C}^∞ decreasing function, so that for $v \in \mathcal{C}_0^\infty$, $R(\lambda, A)v \in \mathcal{C}_0^\infty$. Moreover, a simple computation yields, for all $\lambda > \gamma$,

$$|R(\lambda, A)v(y)| \leq \sup_{(y, \infty)} |v(z)| \frac{1}{\lambda - \gamma} \leq \|v\|_\infty \frac{1}{\lambda - \gamma}.$$

Then

$$\|(P(\varphi R(\lambda, A)))v\|_\infty \leq \|v\|_\infty \frac{\overline{\varphi}}{\lambda - \gamma}$$

and

$$\|(P(\varphi R(\lambda, A)))^n v\|_\infty \leq \|v\|_\infty \left(\frac{\overline{\varphi}}{\lambda - \gamma}\right)^n$$

so that convergence in equation (59) holds in \mathcal{C}^∞ and \mathcal{C}_0^∞ is invariant for $R(\lambda, C)$. The second condition in Proposition 7 follows then by the previous calculations. Finally, because $R(\lambda, C) = (\lambda - C)^{-1}$, to show that $R(\lambda, C)\mathcal{C}_0^\infty$ is dense in \mathcal{C}_0^∞ , it is enough to show that

$$(\lambda - C)\mathcal{C}_0^\infty \subseteq \mathcal{C}_0^\infty.$$

According to Equation (63) and hypothesis (H1) – (H3), this is true. \square

The main result given below shows that when γ_1 is large enough, $u_0(t, y)$ as defined above gives an approximate solution of (49) or (51).

THEOREM 9. *Assume (H1)-(H3) hold. Let $u(0, x, y) \in (\mathcal{C}^\infty)^2 \cap L^1$. For any $\gamma_1 > 0$, let $u(t, x, y; \gamma_1)$ be the associated solution of (48), and define*

$$(64) \quad u_0(t, y; \gamma_1) = \int_0^\infty u(t, x, y; \gamma_1) dx.$$

- (1) *Under the scaling (S1), when $\gamma_1 \rightarrow \infty$, $u_0(t, y; \gamma_1)$ approaches the solution of (49).*
- (2) *Under the scaling (S2) or (S3), when $\gamma_1 \rightarrow \infty$, $u_0(t, y; \gamma_1)$ approaches the solution of (51) with \bar{h} defined by (52).*

In all cases, convergence holds in \mathcal{C}_0^∞ , uniformly in time on any bounded time interval.

PROOF. Throughout the proof, we omit γ_1 in the solution $u(t, x, y; \gamma_1)$ and in the marginal density $u_0(t, y; \gamma_1)$, and keep in mind that they depend on the parameter γ_1 through equation (48).

First, let

$$(65) \quad u_n(t, y) = \int_0^{+\infty} x^n u(t, x, y) dx, \quad n = 0, 1, \dots$$

which are well defined from the previous discussion.

From (48) and (58), we have

$$(66) \quad \begin{aligned} \frac{\partial u_n}{\partial t} = & -n\gamma_1 u_n - \lambda_2 \frac{\partial u_{n+1}}{\partial y} + \gamma_2 \frac{\partial(yu_n)}{\partial y} \\ & + \int_0^\infty \int_0^x \varphi(y) x^n u(t, z, y) h(x - z) dz - \varphi(y) u_n. \end{aligned}$$

Since

$$\int_0^\infty \int_0^x \varphi(y) x^n u(t, z, y) h(x - z) dz = \sum_{j=0}^n \binom{n}{j} \varphi(y) u_{n-j} \mathbb{E}^j h,$$

where

$$\mathbb{E}^j h = \int_0^\infty x^j h(x) dx,$$

we have

$$(67) \quad \frac{\partial u_n}{\partial t} = -n\gamma_1 u_n - \lambda_2 \frac{\partial u_{n+1}}{\partial y} + \gamma_2 \frac{\partial(yu_n)}{\partial y} + \varphi(y) \sum_{j=1}^n \binom{n}{j} u_{n-j} \mathbb{E}^j h.$$

In particular, when $n = 0$,

$$(68) \quad \frac{\partial u_0}{\partial t} = -\lambda_2 \frac{\partial u_1}{\partial y} + \gamma_2 \frac{\partial(yu_0)}{\partial y}.$$

When $n \geq 1$, we have

$$(69) \quad \frac{1}{\gamma_1} \frac{\partial u_n}{\partial t} = -nu_n - \frac{\lambda_2}{\gamma_1} \frac{\partial u_{n+1}}{\partial y} + \frac{\gamma_2}{\gamma_1} \frac{\partial(yu_n)}{\partial y} + \frac{1}{\gamma_1} \varphi(y) \sum_{j=1}^n \binom{n}{j} u_{n-j} \mathbb{E}^j h.$$

When $n \geq 1$ and $\gamma_1 \rightarrow \infty$, we apply the quasi-equilibrium assumption to (69) by assuming

$$\frac{1}{\gamma_1} \frac{\partial u_n}{\partial t} \approx 0$$

when $t \geq t_0 > 0$, and hence

$$(70) \quad u_n = -\frac{\lambda_2}{n\gamma_1} \frac{\partial u_{n+1}}{\partial y} + \frac{\gamma_2}{n\gamma_1} \frac{\partial(yu_n)}{\partial y} + \frac{1}{n\gamma_1} \varphi(y) \sum_{j=1}^n \binom{n}{j} u_{n-j} \mathbb{E}^j h, \quad (\forall n \geq 1).$$

Now, we are ready to prove the results for the three different scalings. (S1). For the scaling (S1), $\varphi(y) \sim \gamma_1$, and we have

$$u_n = \frac{1}{n} \frac{\varphi(y)}{\gamma_1} \sum_{j=1}^n \binom{n}{j} u_{n-j} \mathbb{E}^j h + O\left(\frac{1}{\gamma_1}\right),$$

so

$$(71) \quad u_1 = \frac{b\varphi(y)}{\gamma_1} u_0 + O\left(\frac{1}{\gamma_1}\right),$$

where $b = \mathbb{E}h$. Substituting (71) into (68), we obtain

$$(72) \quad \frac{\partial u_0}{\partial t} = \frac{\partial}{\partial y} [\gamma_2 y u_0 - \lambda_2 \psi(y) u_0] + O\left(\frac{1}{\gamma_1}\right)$$

with $\psi(y) = b\varphi(y)/\gamma_1$. Finally, note that

$$u_0(T, y) = \int_0^T \frac{\partial u_0(t, y)}{\partial t} dt,$$

so (1) follows and convergence holds in \mathcal{C}_0^∞ , uniformly in time on any bounded time interval.

(S2). We assume the scaling (S2) so $h(\Delta X) \rightarrow \frac{1}{\gamma_1} h(\frac{\Delta X}{\gamma_1})$ as $\gamma_1 \rightarrow \infty$, and the re-scaled j^{th} moment

$$b_j = \gamma_1^{-j} \mathbb{E}^j h$$

is independent of γ_1 . Hence, from (70) and Proposition 6, we have

$$\begin{aligned} \gamma_1^{-(n-1)} u_n &= -\frac{\lambda_2}{n} \frac{\partial(\gamma_1^{-n} u_{n+1})}{\partial y} + \frac{\gamma_2}{n\gamma_1} \frac{\partial(y\gamma_1^{-(n-1)} u_n)}{\partial y} + \frac{1}{n} \varphi(y) u_0 b_n \\ &\quad + \frac{1}{n\gamma_1} \varphi(y) \sum_{j=1}^{n-1} \binom{n}{j} \gamma_1^{-(n-j-1)} u_{n-j} b_j \\ &= \frac{1}{n} b_n \varphi(y) u_0 - \frac{\lambda_2}{n} \frac{\partial(\gamma_1^{-n} u_{n+1})}{\partial y} + O\left(\frac{1}{\gamma_1}\right). \end{aligned}$$

Therefore,

$$\begin{aligned} u_1 &= b_1 \varphi(y) u_0 - \lambda_2 \frac{\partial}{\partial y} [\gamma_1^{-1} u_2] + O\left(\frac{1}{\gamma_1}\right) \\ &= b_1 \varphi(y) u_0 - \lambda_2 \frac{\partial}{\partial y} \left[\frac{1}{2} b_2 \varphi(y) u_0 - \frac{\lambda_2}{2} \frac{\partial(\gamma_1^{-2} u_3)}{\partial y} \right] + O\left(\frac{1}{\gamma_1}\right) \\ &= b_1 \varphi(y) u_0 - b_2 \frac{\lambda_2}{2!} \frac{\partial}{\partial y} (\varphi(y) u_0) \\ &\quad + \frac{\lambda_2^2}{2!} \frac{\partial^2}{\partial y^2} \left[\frac{1}{3} b_3 \varphi(y) u_0 - \frac{\lambda_2}{3} \frac{\partial(\gamma_1^{-3} u_4)}{\partial y} \right] + O\left(\frac{1}{\gamma_1}\right) \\ &\dots\dots\dots \\ &= \sum_{k=0}^{\infty} \frac{(-\lambda_2)^k}{(k+1)!} b_{k+1} \frac{\partial^k}{\partial y^k} (\varphi(y) u_0) + O\left(\frac{1}{\gamma_1}\right). \end{aligned}$$

Thus, when $\gamma_1 \rightarrow \infty$, we have

$$\begin{aligned}
-\lambda_2 \frac{\partial u_1}{\partial y} &\rightarrow \sum_{k=1}^{\infty} \frac{(-\lambda_2)^k}{k!} (\gamma_1^{-k} \mathbb{E}^k h) \frac{\partial^k}{\partial y^k} (\varphi(y) u_0) \\
&= \sum_{k=1}^{\infty} \frac{1}{k!} \left(-\frac{\lambda_2}{\gamma_1}\right)^k \left(\int_0^{\infty} x^k h(x) dx\right) \frac{\partial^k}{\partial y^k} (\varphi(y) u_0) \\
&= \int_0^{\infty} \bar{h}(x) \left[\sum_{k=1}^{\infty} \frac{1}{k!} (-x)^k \frac{\partial^k}{\partial y^k} (\varphi(y) u_0) \right] dx \\
&= \int_0^{\infty} \bar{h}(x) (\varphi(y-x) u_0(t, y-x) - \varphi(y) u_0(t, y)) dx \\
&= \int_0^{\infty} \bar{h}(x) \varphi(y-x) u_0(t, y-x) dx - \varphi(y) u_0(t, y) \\
&= - \int_y^{-\infty} \bar{h}(y-z) \varphi(z) u_0(t, z) dz - \varphi(y) u_0(t, y) \\
&= \int_0^y \bar{h}(y-z) \varphi(z) u_0(t, z) dz - \varphi(y) u_0(t, y).
\end{aligned}$$

Here we note $\varphi(z) = 0$ when $z < 0$.

Therefore, from (68), when $\gamma_1 \rightarrow 0$, u_0 approaches to the solution of (51), and the desired result follows.

(S3). Now, we consider the case of scaling (S3) so λ_2/γ_1 is independent of γ_1 . From (70) and Proposition 6, we have

$$\begin{aligned}
u_n &= -\frac{1}{n} \frac{\lambda_2}{\gamma_1} \frac{\partial u_{n+1}}{\partial y} + \frac{\gamma_2}{n\gamma_1} \frac{\partial(y u_n)}{\partial y} + \frac{1}{n\gamma_1} \varphi(y) u_0 \mathbb{E}^n h \\
&\quad + \frac{1}{n\gamma_1} \varphi(y) \sum_{j=1}^{n-1} \binom{n}{j} u_{n-j} \mathbb{E}^j h \\
&= \frac{1}{n\gamma_1} \varphi(y) u_0 \mathbb{E}^n h - \frac{1}{n} \frac{\lambda_2}{\gamma_1} \frac{\partial u_{n+1}}{\partial y} + O\left(\frac{1}{\gamma_1^2}\right).
\end{aligned}$$

Therefore,

$$\begin{aligned}
u_1 &= \frac{1}{\gamma_1} \varphi(y) u_0 \mathbb{E}^1 h - \frac{\lambda_2}{\gamma_1} \frac{\partial}{\partial y} u_2 + O\left(\frac{1}{\gamma_1^2}\right) \\
&= \frac{1}{\gamma_1} \varphi(y) u_0 \mathbb{E}^1 h - \frac{\lambda_2}{\gamma_1} \frac{\partial}{\partial y} \left[\frac{1}{2\gamma_1} \varphi(y) u_0 \mathbb{E}^2 h - \frac{1}{2} \frac{\lambda_2}{\gamma_1} \frac{\partial}{\partial y} u_3 \right] + O\left(\frac{1}{\gamma_1^2}\right) \\
&= \frac{1}{\gamma_1} \varphi(y) u_0 \mathbb{E}^1 h - \frac{1}{2!} \frac{\lambda_2}{\gamma_1^2} \mathbb{E}^2 h \frac{\partial}{\partial y} [\varphi(y) u_0] \\
&\quad + \frac{1}{2!} \left(\frac{\lambda_2}{\gamma_1}\right)^2 \frac{\partial}{\partial y} \left[\frac{1}{3\gamma_1} \varphi(y) u_0 \mathbb{E}^3 h - \frac{1}{3} \frac{\lambda_2}{\gamma_1} \frac{\partial}{\partial y} u_4 \right] + O\left(\frac{1}{\gamma_1^2}\right) \\
&\dots \dots \dots \\
&= -\frac{1}{\lambda_2} \sum_{k=1}^{\infty} \frac{1}{k!} \left(-\frac{\lambda_2}{\gamma_1}\right)^k \mathbb{E}^k h \frac{\partial^{k-1}}{\partial y^{k-1}} [\varphi(y) u_0] + O\left(\frac{1}{\gamma_1^2}\right).
\end{aligned}$$

Thus, when $\gamma_1 \rightarrow \infty$, in a manner similar to the above argument, we have

$$\begin{aligned}
-\lambda_2 \frac{\partial u_1}{\partial y} &\rightarrow \sum_{k=1}^{\infty} \frac{1}{k!} \left(-\frac{\lambda_2}{\gamma_1}\right)^k \mathbb{E}^k h \frac{\partial^k}{\partial y^k} [\varphi(y) u_0] \\
&= \int_0^y \bar{h}(y-z) \varphi(z) u_0(t, z) dz - \varphi(y) u_0(t, y),
\end{aligned}$$

and the result follows. \square

3.3. Result on generators. In this approach, we use a result on semi-groups and the convergence of generators to prove convergence in law for the stochastic process. The main result is

THEOREM 10. *Let $Y(t)$ denote the stochastic process determined by the solution of (1)-(2).*

1. *Assume the scaling (S1), i.e., $\varphi(y)/\gamma_1 = \psi(y)$ as $\gamma_1 \rightarrow \infty$ and h, λ_2 remain unchanged. Then $Y(t)$ converges in law towards the deterministic solution of the ordinary differential equation*

$$(73) \quad \frac{dY}{dt} = -\gamma_2 Y + \lambda_2 b\psi(Y), \quad Y(0) = Y_0 \geq 0.$$

2. *Assume the scaling (S2), i.e., $h \rightarrow \frac{1}{\gamma_1} h(\frac{\Delta X}{\gamma_1})$ as $\gamma_1 \rightarrow \infty$ and φ, λ_2 remain unchanged. Then $Y(t)$ converges in law towards the solution of the stochastic differential equation*

$$(74) \quad \frac{dY}{dt} = -\gamma_2 Y + \dot{N}(\bar{h}, \varphi(Y)), \quad Y(0) = Y_0 \geq 0$$

with $\bar{h}(\Delta Y) = \lambda_2^{-1} h(\lambda_2^{-1} \Delta Y)$ the density function in the jump size ΔY .

3. Assume the scaling (S3), i.e., $\lambda_2/\gamma_1 \rightarrow \bar{\lambda}_2$ as $\gamma_1 \rightarrow \infty$ and h, φ remain unchanged. Then $Y(t)$ converges in law towards the solution of the stochastic differential equation

$$(75) \quad \frac{dY}{dt} = -\gamma_2 Y + \dot{N}(\bar{h}, \varphi(Y)), \quad Y(0) = Y_0 \geq 0$$

with $\bar{h}(\Delta Y) = \bar{\lambda}_2^{-1} h(\bar{\lambda}_2^{-1} \Delta Y)$ the density function in the jump size ΔY .

The proof of the theorem is based on the following result.

KURT'S THEOREM. (*Ethier and Kurtz, 2005, Theorem 2.5, page 167*)
Let X and X^n be Feller processes in $(0, \infty)$ with semigroups (T_t) and $(T_{n,t})$ respectively. Then the following two conditions are equivalent

- $T_{n,t}f \rightarrow T_t f$ for every $f \in \mathcal{C}_0$, uniformly in time for bounded $t > 0$.
- If $X^n(0) \rightarrow X(0)$ in distribution, then $X^n \rightarrow X$ in distribution.

where \mathcal{C}_0 is the class of continuous functions with $f(y) \rightarrow 0$ as $y \rightarrow \infty$.

Here a Feller process is a Markov process with a probability transition function associated with a Feller semigroup as given below. Let $X_t : (0, \infty) \rightarrow \mathbb{R}^+$ be a Markov process, and consider a semigroup defined on the space of bounded continuous functions given by

$$T_t f(x) = \mathbb{E}_x f(X_t),$$

the expectation of the function f along the stochastic process with given starting value at x . By definition, a Markov process is Feller if

- $T_t(\mathcal{C}_0) \subseteq \mathcal{C}_0$
- For every $f \in \mathcal{C}_0$, $\|T_t f\|_\infty \leq \|f\|_\infty$
- For every $f \in \mathcal{C}_0$, $\lim_{t \rightarrow 0} \|T_t f - f\|_\infty = 0$

i.e. if T_t is a strongly continuous contracting semigroup on \mathcal{C}_0 .

LEMMA 11. Assume hypotheses (5)-(6). Then the stochastic processes defined by equations (1)-(2) and by equation (12) are Feller processes.

PROOF. We give a detailed proof for the case of equation (12). The contraction property is immediate. Note that the solution of the deterministic dynamical system, starting at y at $t = 0$, is given by

$$\phi(t, y) = y e^{-\gamma_2 t}$$

Then if $f \in \mathcal{C}_0$ and T_1 is the first jump time, set

$$H(t, y) = \mathbb{P}(T_1 > t) = \exp \left(- \int_0^t \varphi(\phi(s, y)) ds \right),$$

so for any bounded continuous function f we have the following representation for the semigroup P_t (Crudu et al., 2012, Equation 28, page 26)

$$(76) \quad \begin{aligned} T_t f(y) &= f(\phi(t, y)) H(t, y) \\ &+ \int_0^t \int_{\phi(s, y)}^\infty T_{t-s} f(z) h(z - \phi(s, y)) \varphi(\phi(s, y)) H(s, y) dz ds. \end{aligned}$$

For any $y_1 > y_2$, then $\phi(t, y_1) > \phi(t, y_2)$ and

$$\begin{aligned} |T_t f(y_1) - T_t f(y_2)| &\leq |f(\phi(t, y_1)) H(t, y_1) - f(\phi(t, y_2)) H(t, y_2)| \\ &+ \left| \int_0^t \int_{\phi(s, y_1)}^\infty T_{t-s} f(z) \left[h(z - \phi(s, y_1)) \varphi(\phi(s, y_1)) H(s, y_1) \right. \right. \\ &\quad \left. \left. - h(z - \phi(s, y_1)) \varphi(\phi(s, y_1)) H(s, y_1) \right] dz ds \right| \\ &+ \left| \int_0^t \int_{\phi(s, y_2)}^{\phi(s, y_1)} T_{t-s} f(z) h(z - \phi(s, y_2)) \varphi(\phi(s, y_2)) H(s, y_2) dz ds \right| \end{aligned}$$

so that

$$\begin{aligned} |T_t f(y_1) - T_t f(y_2)| &\leq |f(\phi(t, y_1)) H(t, y_1) - f(\phi(t, y_2)) H(t, y_2)| \\ &+ \|f\|_\infty \int_0^t \int_{\phi(s, y_1)}^\infty |h(z - \phi(s, y_1)) \varphi(\phi(s, y_1)) H(s, y_1) \\ &\quad - h(z - \phi(s, y_1)) \varphi(\phi(s, y_1)) H(s, y_1)| dz ds \\ &+ \|f\|_\infty \int_0^t \int_{\phi(s, y_2)}^{\phi(s, y_1)} |h(z - \phi(s, y_2)) \varphi(\phi(s, y_2)) H(s, y_2)| dz ds, \end{aligned}$$

which goes to 0 as $|y_2 - y_1| \rightarrow 0$ by dominated convergence and assumptions (5)-(6). Thus $y \mapsto T_t f(y)$ is continuous. Because the jumps in Y are positive, for all $t \geq 0$ and $y \geq 0$,

$$Y_t^y \geq \phi(t, y).$$

Then, if $f(y) \rightarrow 0$ as $y \rightarrow \infty$, we also have (for fixed t) $f(Y_t^y) \rightarrow 0$ so that $T_t(\mathcal{C}_0) \subseteq \mathcal{C}_0$.

Finally, take $f \in \mathcal{C}_c$ a continuous function with compact support. From equation (76), for any $y \geq 0$,

$$\begin{aligned} |T_t f(y) - f(y)| &\leq |f(\phi(t, y)) - f(y)| H(t, y) + |f(y)| (1 - H(t, y)) \\ &\quad + \int_0^t \int_{\phi(s, y)}^\infty |T_{t-s} f(z) h(z - \phi(s, y)) \varphi(\phi(s, y)) H(s, y)| dz ds \\ &\leq |f(\phi(t, y)) - f(y)| + \|f\|_\infty (1 - \exp(-\overline{\varphi}t)) \\ &\quad + \|f\|_\infty \overline{\varphi} \int_0^t \exp(-\underline{\varphi}s) ds. \end{aligned}$$

Now because $f \in \mathcal{C}_c$, $|f(\phi(t, y)) - f(y)| \rightarrow 0$ as $t \rightarrow 0$, uniformly in y , and

$$\|T_t f(y) - f(y)\|_\infty \rightarrow 0.$$

The theorem is proved using the density of \mathcal{C}_c in \mathcal{C}_0 and the contraction property of T_t . \square

In Kurt's Theorem, by the continuity of the semigroup, to prove the convergence of the stochastic processes it is enough to show the convergence of the semigroup on a dense subset of \mathcal{C}_0 . We therefore restrict our discussion to \mathcal{C}_0^∞ , the class of all infinitely differentiable functions such that f and all its derivatives belong to \mathcal{C}_0 .

To apply this theorem, we show the convergence of the semigroup through their time-derivative formula, which is provided by Dynkin's formula (see Section 3.1).

Proof of Theorem 10. In (43), choose a special test function $g(x, y) = x^n f(y)$ with $f \in \mathcal{C}_0^\infty$. Then we have

$$\begin{aligned} \frac{d}{dt} \mathbb{E}[X_t^n f(Y_t)] &= -\gamma_1 n \mathbb{E}[X_t^n f(Y_t)] - \gamma_2 \mathbb{E}[X_t^n Y_t f'(Y_t)] \\ &\quad + \lambda_2 \mathbb{E}[X_t^{n+1} f'(Y_t)] \\ (77) \quad &\quad + \mathbb{E} \left[\varphi(Y_t) f(Y_t) \left(\int_{X_t}^\infty h(z - X_t) z^n dz - X_t^n \right) \right]. \end{aligned}$$

Since

$$\begin{aligned} \int_x^\infty h(z - x) z^n dz - x^n &= \int_x^\infty h(z - x) (z - x + x)^n dz - x^n \\ &= \int_x^\infty h(z - x) \sum_{k=0}^n \binom{n}{k} (z - x)^k x^{n-k} dz - x^n \\ &= \sum_{k=1}^n \binom{n}{k} x^{n-k} \mathbb{E}^k[h], \end{aligned}$$

where $\mathbb{E}^k[h]$ is the k^{th} moment of h , equation (77) becomes

$$\begin{aligned}
 \frac{d}{dt}\mathbb{E}[X_t^n f(Y_t)] &= -\gamma_1 n \mathbb{E}[X_t^n f(Y_t)] - \gamma_2 \mathbb{E}[X_t^n Y_t f'(Y_t)] \\
 &\quad + \lambda_2 \mathbb{E}[X_t^{n+1} f'(Y_t)] \\
 (78) \quad &\quad + \sum_{k=1}^n \binom{n}{k} \mathbb{E}^k[h] \mathbb{E}[\varphi(Y_t) f(Y_t) X_t^{n-k}].
 \end{aligned}$$

Because f is bounded, that each term in equation (78) is dominated by either $C_f \nu_k(t)$ or $C_f \mu_k(t)$, $k = 0 \cdots n$, where C_f is a constant that depends on f . Now, as in Section 3.1, when $n \geq 1$, the quantity $\mathbb{E}[X_t^n f(Y_t)]$ rapidly approaches its equilibrium value as $\gamma_1 \rightarrow \infty$, so that we can assume

$$\frac{d}{dt}\mathbb{E}[X_t^n f(Y_t)] \simeq 0,$$

which is true for $t \gg \frac{1}{\gamma_1}$. Then, for all $n \geq 1$ and $t > 0$, as $\gamma_1 \rightarrow \infty$,

$$\begin{aligned}
 \mathbb{E}[X_t^n f(Y_t)] &= -\frac{\gamma_2}{\gamma_1 n} \mathbb{E}[X_t^n Y_t f'(Y_t)] \\
 &\quad + \frac{\lambda_2}{\gamma_1 n} \mathbb{E}[X_t^{n+1} f'(Y_t)] \\
 (79) \quad &\quad + \sum_{k=1}^n \binom{n}{k} \frac{\mathbb{E}^k[h]}{\gamma_1 n} \mathbb{E}[\varphi(Y_t) f(Y_t) X_t^{n-k}] + O\left(\frac{1}{\gamma_1}\right).
 \end{aligned}$$

Now, we are ready to prove the results for the three different scaling regimes.

(S1). Letting $n = 1$ in (79), from Proposition 6 and remembering the scaling (S1), for all $t > 0$ and $f \in \mathcal{C}_0^\infty$ and as $\gamma_1 \rightarrow \infty$, we obtain

$$\lambda_2 \mathbb{E}[X_t f'(Y_t)] = \lambda_2 \mathbb{E}[h] \mathbb{E}[\varphi(Y_t) f'(Y_t)] + O\left(\frac{1}{\gamma_1}\right).$$

Now, let $n = 0$ in (78) to obtain

$$(80) \quad \frac{d}{dt}\mathbb{E}[f(Y_t)] = -\gamma_2 \mathbb{E}[Y_t f'(Y_t)] + \lambda_2 \mathbb{E}[h] \mathbb{E}[\varphi(Y_t) f'(Y_t)] + O\left(\frac{1}{\gamma_1}\right).$$

Thus, it is easy to verify that when $\gamma_1 \rightarrow \infty$ and $t \geq \varepsilon > 0$

$$\frac{d}{dt}\mathbb{E}[f(Y_t)] = \mathbb{E}[\mathcal{A}f(Y_t)] + O\left(\frac{1}{\gamma_1}\right),$$

where \mathcal{A} is the generator of the equation (73).

Now consider a sequence $\gamma_{1,n} \rightarrow \infty$, the appropriate scaling φ_n given by (S1), and the associated semi-group T_n . We have, by Dynkin's formula, for all $f \in \mathcal{C}_0^\infty$,

$$T_{n,t}f(y) - T_tf(y) = \int_0^t \frac{d}{ds} (T_{n,s}f(y)) - \frac{d}{ds} (T_sf(y)) ds.$$

Fix any $\varepsilon > 0$. The above calculation shows that, for all $f \in \mathcal{C}_0^\infty$, and $t < T$

$$(81) \quad \|T_{n,t}f(y) - T_tf(y)\| \leq \|T_{n,\varepsilon}f(y) - T_\varepsilon f(y)\| + \frac{C\|f\|}{\gamma_1}(T - \varepsilon),$$

where C is a constant. From the assumption on initial convergence, $\|T_{n,0}f(y) - T_0f(y)\| \rightarrow 0$ as $n \rightarrow \infty$, and by the strong continuity of the semi-group, the first term can then be made arbitrary small. Then the semi-group $T_{n,t}f$ converges uniformly toward T_tf on any bounded time interval. The theorem is proved.

(S2). For the scaling (S2), the proof is similar to that for S1. From Proposition 6, the scaling (S2) and (79) for $n = 1$ gives for all $t > 0$ and $f \in \mathcal{C}_0^\infty$, as $\gamma_1 \rightarrow \infty$,

$$\begin{aligned} \lambda_2 \mathbb{E}[X_t f'(Y_t)] &= -\frac{\lambda_2 \gamma_2}{\gamma_1} \mathbb{E}[X_t Y_t f''(Y_t)] + \frac{\lambda_2^2}{\gamma_1} \mathbb{E}[X_t^2 f''(Y_t)] \\ &\quad + \frac{\lambda_2 \gamma_1 \mathbb{E}[h]}{\gamma_1} \mathbb{E}[\varphi(Y_t) f'(Y_t)] + O\left(\frac{1}{\gamma_1}\right) \end{aligned}$$

Repeating this iteratively, we obtain from (79) and Proposition 6, for all

$t > 0$ and $f \in \mathcal{C}_0^\infty$, as $\gamma_1 \rightarrow \infty$,

$$\begin{aligned}
\lambda_2 \mathbb{E}[X_t f'(Y_t)] &= \lambda_2 \mathbb{E}[h] \mathbb{E}[\varphi(Y_t) f'(Y_t)] \\
&\quad + \frac{\lambda_2^2}{\gamma_1} \left[-\frac{\gamma_2}{2\gamma_1} \mathbb{E}[X_t^2 Y_t f^{(3)}(Y_t)] + \frac{\lambda_2}{2\gamma_1} \mathbb{E}[X_t^3 f^{(3)}(Y_t)] \right. \\
&\quad \left. + 2 \frac{\gamma_1 \mathbb{E}[h]}{2\gamma_1} \mathbb{E}[\varphi(Y_t) f''(Y_t) X_t] \right. \\
&\quad \left. + \frac{\gamma_1^2 \mathbb{E}^2[h]}{2\gamma_1} \mathbb{E}[\varphi(Y_t) f''(Y_t)] \right] + O\left(\frac{1}{\gamma_1}\right) \\
&= \lambda_2 \mathbb{E}[h] \mathbb{E}[\varphi(Y_t) f'(Y_t)] + \frac{\lambda_2^2 \mathbb{E}^2[h]}{2} \mathbb{E}[\varphi(Y_t) f''(Y_t)] \\
&\quad + \frac{\lambda_2^3}{2\gamma_1^2} \mathbb{E}[X_t^3 f^{(3)}(Y_t)] + O\left(\frac{1}{\gamma_1}\right) \\
&\quad \dots \\
(82) \quad &= \sum_{k \geq 1} \frac{\mathbb{E}^k[h] \lambda_2^k}{k!} \mathbb{E}[\varphi(Y_t) f^{(k)}(Y_t)] + O\left(\frac{1}{\gamma_1}\right).
\end{aligned}$$

Since f is analytic, we have

$$\begin{aligned}
\int_y^\infty h(z-y) f(z) dz &= \int_y^\infty h(z-y) f(z-y+y) dz \\
&= \int_y^\infty h(z-y) \sum_{k=0}^\infty \frac{(z-y)^k}{k!} \frac{d^k f}{dy^k}(y) dz \\
&= \sum_{k=0}^\infty \frac{\mathbb{E}^k[h]}{k!} \frac{d^k f}{dy^k}(y) dz.
\end{aligned}$$

Therefore, the last summation term in (82) becomes

$$\sum_{k \geq 1} \frac{\mathbb{E}^k[h] \lambda_2^k}{k!} \mathbb{E}[\varphi(Y_t) f^{(k)}(Y_t)] = \mathbb{E}[\varphi(Y_t) \left(\int_{Y_t}^\infty \bar{h}(z - Y_t) f(z) dz - f(Y_t) \right)],$$

where

$$\bar{h}(y) = \frac{1}{\lambda_2} h\left(\frac{y}{\lambda_2}\right).$$

Now, (78) with $n = 0$ yields

$$\begin{aligned} \frac{d}{dt}\mathbb{E}[f(Y_t)] &= -\gamma_2\mathbb{E}[Y_t f'(Y_t)] \\ &+ \mathbb{E}[\varphi(Y_t)(\int_{Y_t} \bar{h}(z - Y_t)f(z)dz - f(Y_t))] + O\left(\frac{1}{\gamma_1}\right) \\ &= \mathbb{E}[\mathcal{A}f(Y_t)] + O\left(\frac{1}{\gamma_1}\right), \end{aligned}$$

where \mathcal{A} is the generator of equation 74. The rest of the proof is the same as for the scaling (S1) and is omitted.

(S3). Similar to the proof of (2) by simply considering the scaling (S3). The details are omitted.

□

4. From a binary “on-off” model to bursting. In this last section, we show how the convergence of generators used in Section 3.3 is applicable in other cases. In particular, we show how one can obtain a bursting gene expression model of the form (12) from a binary “on-off” model (this has also been proved in Crudu et al. (2012) using different techniques).

The binary on-off model is described by a switching ordinary differential equation (Mackey and Tyran-Kamińska, 2006)

$$(83) \quad \frac{dx}{dt} = -\gamma x + \xi,$$

where ξ is a dichotomous random process, which can take the values 0 or $\lambda > 0$ with switching rates $\alpha(x)$ (from 0 to λ) and $\beta(x)$ (from λ to 0). The evolution equations for densities are

$$(84) \quad \frac{\partial p^0(t, x)}{\partial t} = \frac{\partial(\gamma x p^0(t, x))}{\partial x} - \alpha(x)p^0(x) + \beta(x)p^1(t, x),$$

$$(85) \quad \frac{\partial p^1(t, x)}{\partial t} = \frac{\partial((\gamma x - \lambda)p^1(t, x))}{\partial x} + \alpha(x)p^0(t, x) - \beta(x)p^1(x),$$

where $p^0(t, x)$ means the density function with $\xi(t) = 0$, and $p^1(t, x)$ the density function $p^1(t, x)$ with $\xi(t) = 1$.

This model has been proposed by a number of authors to take into account the stochasticity in the state of the DNA (Peccoud and Ycart, 1995; Innocentini and Hornos, 2007; Lipniacki et al., 2006; Ramos and Hornos, 2007). It is also known that the steady-state solution of this model converges to the steady-state solution of the bursting model when the appropriate scaling is

applied, namely the gene switches very rapidly from on to off ($\beta \rightarrow \infty$), and the synthesis of the gene product is very efficient in the on state ($\lambda \propto \beta$).

The associated semi-group of (83) acts on functions in $\mathcal{C}_0 \times \mathcal{C}_0$, and for any test function $f = (f^0(x), f^1(x)) \in \mathcal{C}_0^\infty \times \mathcal{C}_0^\infty$, the evolution equation of the semi-group is

$$\begin{aligned}
 \frac{d}{dt}T_t(f) &= \frac{d}{dt} \int_0^\infty (p^0(t, x)f^0(x) + p^1(t, x)f^1(x)) dx \\
 (86) \quad &= -\gamma \int_0^\infty xp^0(t, x) \frac{df^0(x)}{dx} dx - \int_0^\infty (\gamma x - \lambda)p^1(t, x) \frac{df^1(x)}{dx} dx \\
 &\quad + \int_0^\infty (\alpha(x)p^0(t, x) - \beta(x)p^1(t, x))(f^1(x) - f^0(x)) dx.
 \end{aligned}$$

For the sake of simplicity we restrict ourselves to the case where $\beta(x)$ remains constant, and consider the scaling

$$(87) \quad \lambda = b\beta.$$

We are going to prove the following result

THEOREM 12. *Assume the scaling (87), and α is bounded in C^1 . For any sequence $\{\beta_n\}$, denote by $X^n(t)$ the stochastic process determined by (83). Suppose $\beta_n \rightarrow \infty$ as $n \rightarrow \infty$, and $\{X^n(0)\}$ converges in distribution to X_0 . Then X^n converges in distribution to the one-dimensional stochastic process $X(t)$ determined by the solution of*

$$(88) \quad \frac{dX}{dt} = -\gamma X + \dot{N}(\bar{h}(\Delta X), \alpha(X)), \quad X(0) = X_0,$$

where \bar{h} is an exponential distribution with mean b .

PROOF. The proof follows along very similar lines as in the previous section. Here we briefly show the calculations.

Consider a function such that $f^0(x) = f^1(x) = f(x)$, and let

$$p(t, x) = p^0(t, x) + p^1(t, x).$$

Then

$$(89) \quad \frac{d}{dt} \int_0^\infty p(t, x)f(x)dx = -\gamma \int_0^\infty xp(t, x) \frac{df}{dx} dx + \lambda \int_0^\infty p^1(t, x) \frac{df}{dx} dx.$$

To calculate the last term, we note for any $g \in \mathcal{C}_0^\infty$,

$$(90) \quad \begin{aligned} \frac{d}{dt} \int_0^\infty p^1(t, x) g(x) dx &= -\gamma \int_0^\infty x p^1(t, x) \frac{dg}{dx} dx + \lambda \int_0^\infty p^1(t, x) \frac{dg}{dx} dx \\ &+ \int_0^\infty \alpha(x) p^0(t, x) g(x) dx - \int_0^\infty \beta p^1(t, x) g(x) dx \end{aligned}$$

When $\beta \rightarrow \infty$, one can make a quasi steady-state approximation to (90), which gives

$$\begin{aligned} \int_0^\infty p^1(t, x) g(x) dx &= -\frac{\gamma}{\beta} \int_0^\infty x p^1(t, x) \frac{dg}{dx} dx + \frac{1}{\beta} \int_0^\infty \alpha(x) p^0(t, x) g(x) dx \\ &+ \frac{\lambda}{\beta} \int_0^\infty p^1(t, x) \frac{dg}{dx} dx + O\left(\frac{1}{\beta}\right). \end{aligned}$$

Iterating the process and writing $g = \frac{df}{dx}$, we have

$$(91) \quad \begin{aligned} \int_0^\infty \lambda p^1(t, x) \frac{df}{dx} dx &= \sum_{i \geq 1} \left(\frac{\lambda}{\beta}\right)^i \int_0^\infty \alpha(x) p(t, x) \frac{d^i f}{dx^i} \\ &- \sum_{i \geq 1} \left(\frac{\lambda}{\beta}\right)^i \int_0^\infty \gamma x p^1(t, x) \frac{d^i f}{dx^i} \\ &- \sum_{i \geq 1} \left(\frac{\lambda}{\beta}\right)^i \int_0^\infty \alpha(x) p^1(t, x) \frac{d^i f}{dx^i} + O\left(\frac{1}{\beta}\right) \end{aligned}$$

The first sum gives the jump kernel

$$\int_0^\infty \alpha(x) p(t, x) \left(\int_x^\infty h(y - x) f(y) dy - f(x) \right) dx,$$

where h is a distribution whose moments are given by

$$\mathbb{E}^i[h] = i! \left(\frac{\lambda}{\beta}\right)^i = i! b^i.$$

The other two summations are dominated by

$$\begin{aligned} &\left| \sum_{i \geq 1} \left(\frac{\lambda}{\beta}\right)^i \int_0^\infty \gamma x p^1(t, x) \frac{d^i f}{dx^i} + \sum_{i \geq 1} \left(\frac{\lambda}{\beta}\right)^i \int_0^\infty \alpha(x) p^1(t, x) \frac{d^i f}{dx^i} \right| \\ &\leq C(f) \int_0^\infty x p^1(t, x) dx. \end{aligned}$$

This last integral goes to 0 as $\beta \rightarrow \infty$ because

$$\begin{aligned} \frac{d}{dt} \int_0^\infty xp^1(t, x) dx &= \int_0^\infty (\alpha(x) - \gamma - \beta) xp^1(t, x) dx + \lambda \int_0^\infty p^1(t, x) dx \\ &\leq (\alpha_\infty - \gamma - \beta) \int_0^\infty xp^1(t, x) dx + \lambda \int_0^\infty p^1(t, x) dx \end{aligned}$$

which allow us to apply the Gronwall's lemma and the fact that $\int_0^\infty p^1(t, x) dx \rightarrow 0$ to conclude the claim is true.

Now, substitute (92) into (89). When $\beta \rightarrow \infty$, we obtain

$$\begin{aligned} \frac{d}{dt} \int_0^\infty p(t, x) f(x) dx &= -\gamma \int_0^\infty xp(t, x) \frac{df}{dx} dx \\ &+ \int_0^\infty \alpha(x) p(t, x) \left(\int_x^\infty h(y - x) f(y) dy - f(x) \right) dx \end{aligned}$$

which is the same as we obtained from the equation (88).

The remainder of the proof proceeds as in the proof of Theorem 10 and the details are omitted. \square

References.

- BALL, K., KURTZ, T. G., POPOVIC, L. and REMPALA, G. (2006). Asymptotic Analysis of Multiscale Approximations to Reaction Networks. *The Annals of Applied Probability* **16** 1925–1961.
- BERGLUND, N. and GENTZ, B. (2006). *Noise-Induced Phenomena in Slow-Fast Dynamical Systems, A Sample-Paths Approach*. Springer.
- BUJORIANU, M. L. and KATOEN, J. P. (2008). Symmetry reduction for stochastic hybrid systems. In *Decision and Control, 2008. CDC 2008. 47th IEEE Conference on DOI - 10.1109/CDC.2008.4739086* 233–238.
- BUJORIANU, M. L. and LYGEROS, J. (2004). General stochastic hybrid systems: modelling and optimal control. In *Decision and Control, 2004. CDC. 43rd IEEE Conference on DOI - 10.1109/CDC.2004.1430320* **2** 1872–1877 Vol.2.
- CRUDU, A., DEBUSSCHE, A., MULLER, A. and RADULESCU, O. (2012). Convergence of stochastic gene networks to hybrid piecewise deterministic processes. To appear in *Annals of Applied Prob.*
- CCERES, M. O. and BUDINI, A. A. (1997). The generalized Ornstein - Uhlenbeck process. *Journal of Physics A: Mathematical and General* **30** 8427-.
- DAVIS, M. H. A. (1984). Piecewise-Deterministic Markov Processes: A General Class of Non-Diffusion Stochastic Models. *Journal of the Royal Statistical Society. Series B (Methodological)* **46** 353–388.
- ELF, J., LI, G.-W. and XIE, X. S. (2007). Probing Transcription Factor Dynamics at the Single-Molecule Level in a Living Cell. *Science* **316** 1191–1194.
- ETHIER, S. N. and KURTZ, T. G. (2005). *Markov Processes: Characterization and Convergence*. Wiley Interscience.
- FENICHEL, N. (1979). Geometric singular perturbation theory for ordinary differential equations. *Journal of Differential Equations* **31** 53–98.

- FRIEDMAN, N., CAI, L. and XIE, X. S. (2006). Linking Stochastic Dynamics to Population Distribution: An Analytical Framework of Gene Expression. *Phys. Rev. Lett.* **97** 168302–.
- GARDINER, C. W. (1985). *Handbook of stochastic methods, for physics, chemistry and the natural sciences*, second ed. *Springer Series in Synergetics* **13**. Springer-Verlag, Berlin. [MR858704 \(87i:00033\)](#)
- GEL'FAND, I. M. and VILENKIN, N. Y. (1964). *Applications of Harmonic Analysis. Generalized Functions* **4**. Academic Press, New York.
- GOLDING, I., PAULSSON, J., ZAWILSKI, S. M. and COX, E. C. (2005). Real-Time Kinetics of Gene Activity in Individual Bacteria. *Cell* **123** 1025–1036.
- GOUTSIAS, J. (2005). Quasiequilibrium approximation of fast reaction kinetics in stochastic biochemical systems. *J. Chem. Phys.* **122** 184102–15.
- HASELTINE, E. L. and RAWLINGS, J. B. (2002). Approximate simulation of coupled fast and slow reactions for stochastic chemical kinetics. *J. Chem. Phys.* **117** 6959–6969.
- HASELTINE, E. L. and RAWLINGS, J. B. (2005). On the origins of approximations for stochastic chemical kinetics. *J. Chem. Phys.* **123** 164115–16.
- HESPANHA, J. P. (2006). Modelling and analysis of stochastic hybrid systems. *Control Theory and Applications, IEE Proceedings - DOI -* **153** 520–535.
- HIDA, T. and SI, S. (2004). *An Innovation approach to random fields, Application of White Noise Theory*. World Scientific Pub. Co.
- HIDA, T. and SI, S. (2008). *Lectures on White Noise Functionals*. World Scientific Pub. Co.
- INNOCENTINI, G. and HORNOS, J. (2007). Modeling stochastic gene expression under repression. *Journal of Mathematical Biology* **55** 413–431.
- JACOD, J. and SHIRYAEV, A. N. (1987). *Limit theorems for stochastic processes*. Springer.
- KANG, H.-W. and KURTZ, T. G. (2012). Separation of time-scales and model reduction for stochastic reaction networks. To appear in *Annals of Applied Prob.*
- LASOTA, A. and MACKEY, M. C. (1985). *Probabilistic properties of deterministic systems*. Cambridge University Press, Cambridge. [MR832868 \(87i:58100\)](#)
- LEI, J. (2009). Stochastic in single gene expression with both intrinsic noise and fluctuation in kinetic parameters. *Journal of Theoretical Biology* **256** 485–492.
- LIPNIAK, T., PASZEK, P., MARCINIAK-CZOCZRA, A., BRASIER, A. R. and KIMMEL, M. (2006). Transcriptional stochasticity in gene expression. *Journal of Theoretical Biology* **238** 348–367.
- MACKEY, M. C. and TYRAN-KAMIŃSKA, M. (2006). Noise and Conditional Entropy Evolution. *Physica A* **365** 360–382.
- MACKEY, M. C. and TYRAN-KAMIŃSKA, M. (2008). Dynamics and density evolution in piecewise deterministic growth processes. *Ann. Polon. Math.* **94** 111–129. . [MR2438853 \(2009h:47072\)](#)
- MACKEY, M. C., TYRAN-KAMINSKA, M. and YVINEC, R. (2011). Molecular distributions in gene regulatory dynamics. *Journal of Theoretical Biology* **274** 84 - 96.
- OZBUDAK, E. M., THATTAI, M., KURTSE, I., GROSSMAN, A. D. and VAN OUDENAARDE, A. (2002). Regulation of noise in the expression of a single gene. *Nat Genet* **31** 69–73.
- PAZY, A. (1992). *Semigroups of Linear Operators and Applications to Partial Differential Equations*. Springer.
- PECCOUD, J. and YCART, B. (1995). Markovian Modeling of Gene-Product Synthesis. *Theoretical Population Biology* **48** 222–234.
- PROKHOROV, Y. V. (1956). Convergence of Random Processes and Limit Theorems in Probability Theory. *Theory Probab. Appl.* **1** 157–214.

- RAJ, A. and VAN OUDENAARDEN, A. (2009). Single-Molecule Approaches to Stochastic Gene Expression. *Annu. Rev. Biophys.* **38** 255–270.
- RAJ, A., PESKIN, C. S., TRANCHINA, D., VARGAS, D. Y. and TYAGI, S. (2006). Stochastic mRNA Synthesis in Mammalian Cells. *PLoS Biol* **4** e309–.
- RAMOS, A. F. and HORNOS, J. E. M. (2007). Symmetry and Stochastic Gene Regulation. *Phys. Rev. Lett.* **99** 108103–.
- RAO, C. V. and ARKIN, A. P. (2003). Stochastic chemical kinetics and the quasi-steady-state assumption: Application to the Gillespie algorithm. *J. Chem. Phys.* **118** 4999–5010.
- RUDIN, W. (1991). *Functional analysis*. McGraw-Hill Scienc.
- SCHAEFER, H. H. (1971). *Topological Vector Spaces*. Springer-Verlag New York Heidelberg Berlin.
- SHAHREZAEI, V. and SWAIN, P. S. (2008). Analytical distributions for stochastic gene expression. *Proc Natl Acad Sci U S A* **105** 17256–17261.
- SUTER, D. M., MOLINA, N., GATFIELD, D., SCHNEIDER, K., SCHIBLER, U. and NAEF, F. (2011). Mammalian Genes Are Transcribed with Widely Different Bursting Kinetics. *Science* **332** 472–474.
- TIKHONOV, A. N. (1952). Systems of differential equations containing small parameters in the derivatives. *Mat. Sb. (N.S.)* **31 (73)** 575–586.
- TYRAN-KAMINSKA, M. (2009). Substochastic semigroups and densities of piecewise deterministic Markov processes. *Journal of Mathematical Analysis and Applications* **357** 385–402.
- XIE, X. S., CHOI, P. J., LI, G.-W., LEE, N. K. and LIA, G. (2008). Single-Molecule Approach to Molecular Biology in Living Bacterial Cells. *Annu. Rev. Biophys.* **37** 417–444.
- ZERON, E. S. and SANTILLÁN, M. (2010). Distributions for negative-feedback-regulated stochastic gene expression: Dimension reduction and numerical solution of the chemical master equation. *Journal of Theoretical Biology* **264** 377–385.

ZHOU PEI-YUAN CENTER FOR APPLIED MATHEMATICS
TSINGHUA UNIVERSITY, BEIJING 100084, CHINA
E-MAIL: jzlei@tsinghua.edu.cn; zgcj08@mails.tsinghua.edu.cn

DEPARTMENTS OF PHYSIOLOGY, PHYSICS & MATHEMATICS
CENTRE FOR APPLIED MATHEMATICS IN BIOSCIENCE & MEDICINE
3655 PROMENADE SIR WILLIAM OSLER
MONTREAL, QC, CANADA, H3G 1Y6
E-MAIL: michael.mackey@mcgill.ca

UNIVERSITÉ DE LYON CNRS UMR 5208 UNIVERSIT LYON 1
INSTITUT CAMILLE JORDAN 43 BLVD. DU 11 NOVEMBRE 1918
F-69622 VILLEURBANNE CEDEX FRANCE
E-MAIL: [yvynec@math.univ-lyon1.fr](mailto:yvinec@math.univ-lyon1.fr)